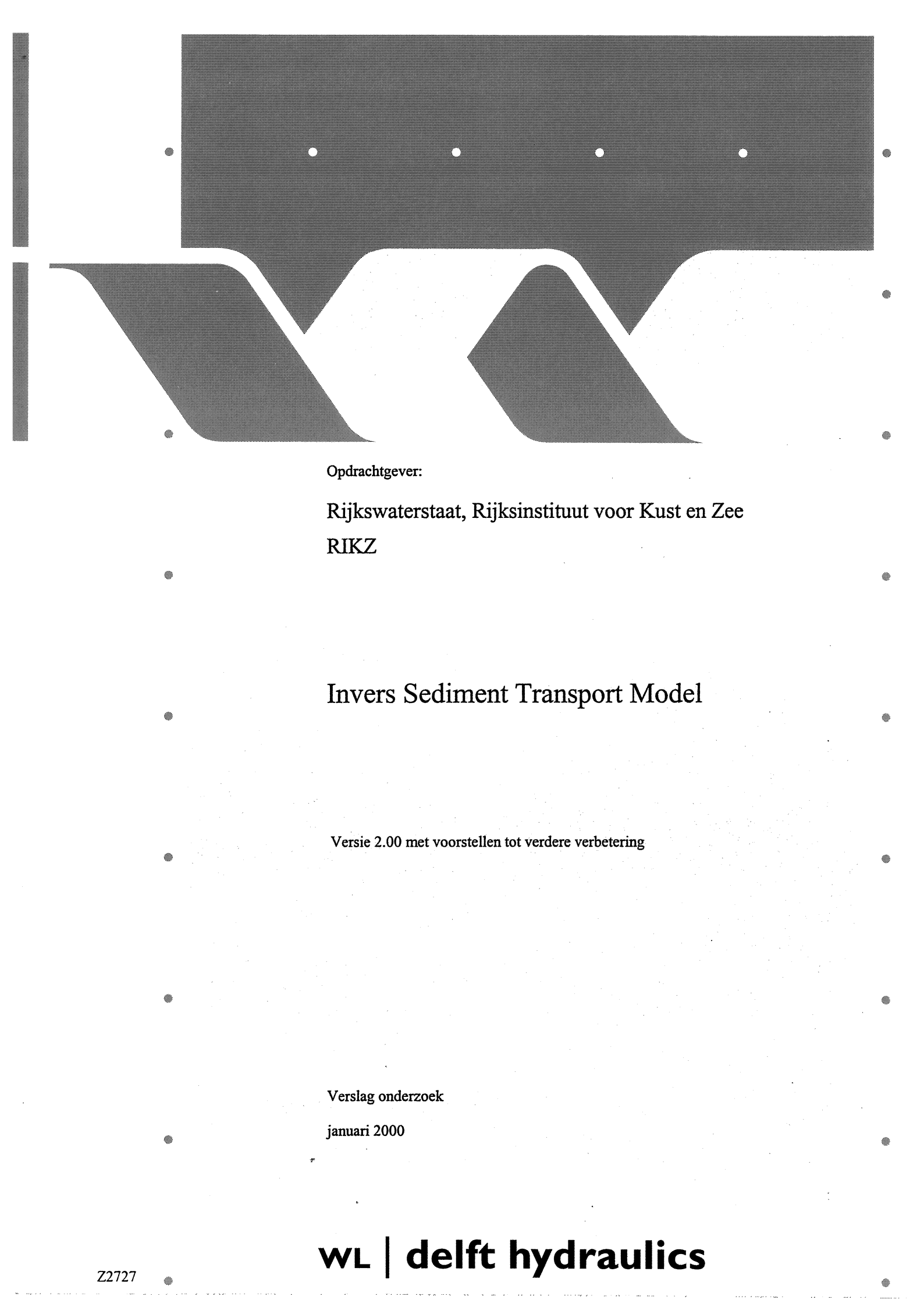




Opdrachtgever:

Rijkswaterstaat, Rijksinstituut voor Kust en Zee
RIKZ

Invers Sediment Transport Model

Versie 2.00 met voorstellen tot verdere verbetering

Verslag onderzoek

januari 2000

Invers Sediment Transport Model

Versie 2.00 met voorstellen tot verdere verbetering

H.F.P. van den Boogaard, J.C. Winterwerp, A.W.J. Koster



wl | delft hydraulics



OPDRACHTGEVER: Rijkswaterstaat, Rijksinstituut voor Kust en Zee, RIKZ

TITEL: Invers Sediment Transport Model. Versie 2.00 met voorstellen tot verdere verbetering.

SAMENVATTING: Met het Invers Sediment Transport Model (ISTM) kunnen diverse bronnen van informatie van sediment-transporten worden geïntegreerd die voor een bepaald gebied beschikbaar zijn. Het model is sinds 1994 door RIKZ voor diverse, grootschalige praktijktoepassingen ingezet. De daarin opgedane ervaringen hebben aanleiding gegeven om te onderzoeken of het huidige modelconcept op de volgende punten verbeterd zou kunnen worden: (1) het inbrengen van meer fysika in het model, (2) de wijze waarop transportrichtingen in het model verwerkt worden, en (3) het conceptuele model en de functie van regularisatie. De resultaten van een onderzoek naar deze verbeteringen zijn vastgelegd in de hier voorliggende rapportage en kunnen als volgt worden samengevat.

Voor het inbrengen van meer *fysische voorkennis* worden de meeste perspectieven verwacht van morfologische voorkennis door een ervaren morfoloog ten aanzien van transportrichtingen door geulen en rondom plaat-geulovergangen.

De mogelijkheden tot het inbrengen van *fysische voorkennis* op basis van *empirische relaties* waarmee afhankelijkheden van bepaalde geometrische karakteristieken van een estuarium zijn vastgelegd, zijn beperkter. De empirische evenwichtsrelaties lijken dan met name het meest geschikt voor data-validatie.

De fysisch-georiënteerde informatie gegenereerd met *initiële procesmodellen* lijkt bruikbaar om het ISTM met 'virtuele metingen' te voeden. Hierbij moet echter wel zorgvuldig worden nagegaan of de uitkomsten consistent zijn met de gemeten informatie.

Voor de bruikbaarheid van zogenaamde *hybride* modellen kunnen dezelfde overwegingen worden gegeven als hierboven voor de initiële procesmodellen. Een belangrijk verschil is echter dat die informatie nu veel meer groottes en richtingen van transporten bevat over langere, morfologische tijdsschalen, en in spatiële zin dan meer op makro-schaal beschikbaar/representatief is.

Voor het in het ISTM inbrengen van meer *data-georiënteerde voorkennis* wordt aanbevolen om aan de puntmetingen zoals die nu aan het ISTM worden aangeboden eveneens geaggregeerde grootheden toe te voegen. Deze kunnen variëren van totale of gemiddelde (grootte of richtingen van) transporten over een raai of een deelgebied, tot zelfs termen in balansvergelijkingen op het niveau van morfologische eenheden.

Aan de hand van een algemene review van *regularisatie* zijn een aantal aanbevelingen afgeleid voor het verbeteren van dit mechanisme. Een van de conclusies is dat het hierin aanbrengen van ruimtelijke variaties belangrijke voordelen kan bieden t.o.v. het uniform toepassen zoals dat tot nu toe moest worden gedaan. Een tweede belangrijke verbetering kan worden bereikt door het regularisatie-mechanisme meer flexibel/generiek op te zetten.

Voor het verbeteren van de wijze waarop in het ISTM informatie van *transportrichtingen* wordt verwerkt is het conceptuele model aangepast t.a.v. de richtingsvergelijkingen. Een belangrijke konsekwentie van deze verbetering is dat het ISTM moest worden uitgebreid naar een niet-lineaire vorm. Dit heeft geleid tot de operationalisatie van versie 2.00 als opvolger van de vorige versie 1.03.

REFERENTIES: Opdrachtverlening overeenkomst RKZ-737, RIKZ/OS 996552, K2000*KOP, 17 augustus 1999
Wijzigingsovereenkomst RKZ-737A, RIKZ/OS 997535, K2000*KOP, 24 december 1999

REV.	AUTEUR	DATUM	OPMERKINGEN	REVIEW	GOEDKEURING
01	H.F.P. van den Boogaard	Nov. 1999		Z.B. Wang	T. Schilperoort
	J.C. Winterwerp	Nov. 1999			
02	H.F.P. van den Boogaard	Jan. 2000			
	J.C. Winterwerp	Jan. 2000			
TREFWOORDEN			INHOUD		STATUS
Invers Sediment Transport Model, Fysische en morfologische systeemkennis, Niet-uniforme Regularisatie, Niet-lineaire richtingsvergelijkingen, Niet-lineair ISTM, Kostenfunctie, Gradiënt methodes			TEKST:	73	☐ VOORLOPIG
			TABELLEN:		☐ CONCEPT
			FIGUREN:	14	☒ DEFINITIEF
			APPENDICES:	2	
PROJECTNUMMER: Z2727					

Inhoud

Lijst van Symbolen

Samenvatting	1
1 Inleiding	1-1
2 Fysische voorkennis	2-1
2.1 Empirische voorkennis	2-1
2.2 Hybride modellen	2-3
2.3 Procesmodellen	2-4
2.4 Aanbevelingen	2-6
3 Data-georiënteerde voorkennis	3-1
4 Regularisatie-vergelijkingen	4-1
4.1 Review achtergrond van regularisatie	4-1
4.2 Regularisatie als filter in het ISTM	4-10
4.3 Overige vormen voor het inbrengen van regularisatie	4-14
4.4 Samenvatting aanbevelingen regularisatie	4-16
5 Opzet van het niet-lineaire ISTM	5-1
5.1 Review van het lineaire ISTM model	5-1
5.2 Definitie van het niet-lineaire ISTM	5-4
6 Transportrichtingen in het niet-lineaire ISTM	6-1
6.1 Optie 1. ISTM model voor strikt parallelle richtingen	6-3
6.2 Optie 2. ISTM model voor parallelle of anti-parallelle richtingen	6-6
7 Discussie en aanbevelingen	7-1
Referenties	Ref.-1

Appendices

A	User manual ISTM, Versie 2.00	A-1
B	Testen van het richtingenformalisme in het niet lineaire ISTM	B-1

Lijst van Symbolen

Symbol **Betekenis**

A_c	Geuldoorsnede onder gemiddeld water (MSL)
P	Getijprisma ter plaatse van A_c
$H_{h,e}$	Evenwichtshoogte t.o.v. MSL van het intertijdegebied onder MSL
$H_{l,e}$	Evenwichtshoogte t.o.v. MSL van het intertijdegebied boven MSL
a_l	Getijamplitude
A_b	Totale oppervlak van een getijbassin
A_{ig}	Totale areaal aan intertijdegebieden
A_{MHW}	Bassinoppervlak onder gemiddelde hoogwater (MHW)
$\vec{T}$	Twee dimensionaal, horizontaal transportvektorveld; $\vec{T}(x, y) = (T_x(x, y), T_y(x, y))^T$
$(T_x)_m^n$	x -component van $\vec{T}$ voor cel (m, n) van het rechthoekige, spatiële numerieke rooster
$(T_y)_m^n$	y -component van $\vec{T}$ voor cel (m, n) van het rechthoekige, spatiële numerieke rooster
$\hat{T}$	Gemeten transportvektor
B	Bergingenveld; $B = B(x, y)$
$\hat{B}$	Gemeten berging
φ	Transportrichting; $\varphi(\vec{T}) = \arctan 2(T_y, T_x)$
$\hat{\varphi}$	Gemeten transportrichting
$\vec{\Theta}$	Vektor die volgens bepaalde nummering over (m, n) -rooster alle $(T_x)_m^n, (T_y)_m^n$ omvat
$\vec{H}(\vec{\Theta})$	Deterministisch deel van het stelsel ISTM-vergelijkingen
$H_n(\vec{\Theta})$	Deterministisch deel van de n^e ISTM-vergelijking
V_n	Ruis die de onzekerheid aangeeft van de n^e ISTM-vergelijking
$\vec{V}$	Ruisvektor die de diverse V_n omvat
σ_n	Spreiding van de ruis V_n die de onzekerheid aangeeft van de n^e ISTM-vergelijking
R	Co-variantiematrix van een verzameling ruizen V_n ; $R = E[\vec{V} \cdot \vec{V}^T], R_{\mu\nu} = E[V_\mu \cdot V_\nu]$
$\vec{Z}$	Rechterlid in lineaire ISTM-vergelijkingen
$E[\cdot]$	Verwachtingswaarde operator
$\tilde{H}(\cdot)$	Met een co-variantiematrix $R^{-\frac{1}{2}}$ geschaalde $\vec{H}(\cdot)$
$\tilde{V}$	Met een co-variantiematrix $R^{-\frac{1}{2}}$ geschaalde ruisvektor $\vec{V}$
$\tilde{Z}$	Met een co-variantiematrix $R^{-\frac{1}{2}}$ geschaalde rechterlid in ISTM vergelijkingen

Symbol	Betekenis
$J(\cdot)$	Kostenfunctie
$\vec{\nabla} J$	Gradiënt van de kostenfunctie $J(\cdot)$
$L(\cdot)$	Kostenfunctie $J(\cdot)$ uitgebreid met regularisatietermen
Δx	celgrootte voor de diskretisatie in de x -richting
Δy	celgrootte voor de diskretisatie in de y -richting
$\ \vec{r}\ $	Standaard (Euklidische) norm van een vektor $\vec{r}$
$\ \vec{r}\ _{\Sigma^{-1}}$	Norm van een vektor $\vec{r}$ bij een niet isotrope afstandsdefinitie
ν	spatiële hoekfrequentie
P_{Θ}	Co-variantiematrix van een oplossing $\vec{\Theta}$ van de ISTM vergelijkingen

Samenvatting

In 1993-1994 is door WL | Delft Hydraulics in opdracht van Rijkswaterstaat/RIKZ het Invers Sediment Transport Model (ISTM) ontwikkeld voor het invers modelleren van twee dimensionale horizontale sedimenttransporten. Met dit ISTM kunnen diverse bronnen van informatie van sedimenttransporten worden geïntegreerd die voor een bepaald gebied beschikbaar zijn. Het resultaat van een ISTM-berekening is dan een transportenveld dat voor het beschouwde gebied het over een tijdsinterval gemiddelde sedimenttransport weergeeft dat optimaal past bij de gegeven informatie.

Het lineaire ISTM (versie 1.03) is sindsdien door RIKZ voor diverse, grootschalige praktijktoepassingen ingezet. De daarin opgedane ervaringen hebben aanleiding gegeven om in nader detail te onderzoeken of het modelconcept op een aantal punten verbeterd zou kunnen worden. Hierbij werd expliciet aan de volgende drie aspecten gedacht:

1. Het inbrengen van meer fysika in het model
2. De (her)definitie van de wijze waarop transportrichtingen in het model verwerkt worden
3. De functie van regularisatie in het ISTM

De hier voorliggende rapportage bevat de resultaten van een door WL | Delft Hydraulics in opdracht van Rijkswaterstaat/RIKZ uitgevoerde studie naar de mogelijkheden om op voornoemde drie punten het ISTM te verbeteren. Hierbij moet worden aangetekend dat deze studie in overmaat een oriëntatie/onderzoekskarakter heeft gehad in de zin dat op basis van materiedeskundigheid en literatuuronderzoek een inventarisatie van potentiële verbeteringen is uitgevoerd. Het konkrete resultaat van de studie is daarmee bevat in diverse aanbevelingen voor de beoogde verbeteringen. Deze verbeteringen zijn op dit moment alleen voor het 2^e aandachtspunt, namelijk aanpassing van het model voor het verwerken van richtingsinformatie geïmplementeerd en heeft geleid tot een niet-lineair ISTM (versie 2.0). T.a.v. de aanbevelingen voor de overige twee aspecten is volstaan met de bijbehorende mathematische formuleringen en deze zijn nog niet daadwerkelijk geïmplementeerd, en dus ook niet getoetst aan de hand van één of meer praktijktoepassingen. Zo'n implementatie in het niet-lineaire ISTM (versie 2.0) en toetsing aan de hand van praktijktoepassingen vormt een belangrijk onderwerp voor verder onderzoek.

In hoofdlijnen kunnen de resultaten van dit onderzoek als volgt worden samengevat.

Voor het inbrengen van meer *fysische voorkennis* in het ISTM worden de meeste perspectieven verwacht van morfologische voorkennis door een ervaren morfoloog ten aanzien van transportrichtingen door geulen en rondom plaat-geulovergangen. Deze informatie kan als extra metingen, dan wel als regularisatie aan het ISTM worden toegevoegd. Deze fysische voorkennis wordt door het ISTM dan geïntegreerd met de data van de overige informatiebronnen. Bij de analyse/interpretatie van het uiteindelijke resultaat moet er dan wel altijd voor worden gewaakt dat de daarin teruggevonden aspecten van de systeemkennis dan geheel of voor een groot gedeelte het gevolg kunnen zijn van die toegevoegde ISTM-invoer. Het waarheidsgehalte van de betreffende uitvoer is daarmee dus niet altijd groter geworden dan (of wordt bevestigd door) diens 'a priori' status, d.w.z. ten tijde van het feitelijke toevoegen van de fysische voorkennis aan de invoer. Uiteraard kan

mede aan de hand van de uitvoer in sommige gevallen wel worden nagegaan óf, en in welke mate, de toegevoegde fysische voorkennis consistent is met de overige input data. Bij een eventuele consistentie is de 'waarheid' van die voorkennis nog steeds niet in strikte zin 'bewezen' maar kan misschien wel in meer of mindere mate aannemelijk worden gemaakt, om aldus het toevoegen van die voorkennis aan de ISTM invoer te rechtvaardigen.

In meer konkrete zin, en redenerend vanuit (de praktijk en doelstelling bij) toepassingen, moeten in het bijzonder de tijd- en ruimteschalen en/of het wel of niet voldaan zijn aan evenwichtskondities worden betrokken bij het hierboven aanbevolen consistent inbrengen van fysische kennis, en het daarna voorzichtig/kritisch interpreteren van de ISTM-resultaten. Deze kondities en hun consistentie bepalen in sterke mate de meerwaarde van de toegevoegde kennis, en geldt zowel voor de hierboven genoemde fysische voorkennis als voor de overige hieronder genoemde additionele informatie-bronnen.

De mogelijkheden tot het inbrengen van *fysische voorkennis* op basis van *empirische relaties* waarmee afhankelijkheden van bepaalde geometrische karakteristieken van een estuarium zijn vastgelegd, zijn beperkter. Bij die geometrische karakteristieken en hun relaties moet met name worden gedacht aan de totale arealen van intergetijdegebieden met hun relatie tot het totale bassinoppervlak dat onder gemiddeld hoogwater is gelegen. De hieruit af te leiden, en aan het ISTM toe te voegen vergelijkingen leveren geaggregeerde transporteninformatie. Geaggregeerd betekent dat deze informatie is samengesteld uit over deelgebieden geïntegreerde transportgroottes of transportrichtingen, en dus niet zozeer geassocieerd zijn met één of een relatief klein aantal naburige punten van het numerieke rooster. Merk hierbij op dat in het huidige ISTM alle invoergegevens wel als puntbron worden verwerkt. Een belangrijke beperking van de bedoelde empirische relaties is dat deze meer informatie leveren over richtingen dan over absolute groottes van transporten.

Voorgesteld wordt om de genoemde empirische evenwichtsrelaties met name te gebruiken voor het toetsen van de consistentie van deze informatie met de overige invoer data (meetgegevens in strikte zin). Aan de hand van zo'n controle kunnen dan gerichte selecties - of voorbewerkingen van de ISTM-invoer worden gepleegd.

Op de tweede plaats lijkt de fysisch-georiënteerde informatie gegenereerd met *initiële procesmodellen* bruikbaar om het ISTM met 'virtuele metingen' te voeden. Het grote voordeel van deze aanpak is dat de toegevoegde informatie op elke ruimtelijke schaal beschikbaar is, en de bijbehorende massabalans intern consistent is. Bovendien zijn de vergelijkingen om deze informatie aan het ISTM te koppelen al grotendeels in de huidige versie bevat. Desalniettemin moet bij deze aanpak wel de nodige voorzichtigheid worden betracht. Vanwege de beperkte nauwkeurigheid van de gebruikte sedimenttransport-formules, alsmede onzekerheden of de fysische kondities waaronder die initiële procesmodellen toegepast mogen worden compatibel zijn met de feitelijke toestand, moet zorgvuldig worden nagegaan of de uitkomsten van zo'n initieel procesmodel consistent zijn met de gemeten informatie. Dit vergt dus weer een zorgvuldige controle en/of selectie en/of voorbewerkingen van de op deze manier gegenereerde additionele data.

Als derde bruikbare bron voor het genereren van additionele fysische informatie voor het ISTM kunnen de resultaten van hybride modellen worden genoemd. Hybride modellen worden gekenmerkt door de integratie van empirische kennis van een systeem met conceptuele numerieke modellen voor het berekenen van de waterbeweging en aspecten van het sedimenttransport. Zij vormen daarmee een tussenvorm van de hiervoor genoemde volledig empirische relaties enerzijds, en conceptuele procesmodellen anderzijds. Voor de bruikbaarheid van hybride modellen, en de manier waarop hun resultaten naar een ISTM

omgeving kunnen worden overgeheveld kunnen dezelfde overwegingen worden gegeven als hierboven voor de initiële procesmodellen. Een belangrijk verschil is echter dat die informatie nu veel meer groottes en richtingen van transporten bevat over langere, morfologische tijdsschalen, en in spatiële zin meer op makro-schaal beschikbaar/representatief is dan in de vorm van puntmetingen.

Naast de hierboven genoemde bronnen van *fysische voorkennis* (waarbij uitgegaan wordt van generieke fysische eigenschappen van sedimenttransporten en/of morfologische processen) zijn de perspectieven van het in het ISTM inbrengen van meer *data-georiënteerde voorkennis* (waarbij meer vanuit data en gebiedskennis wordt uitgegaan, en daarmee specifiek voor de feitelijke case) nagegaan. Hierbij wordt aanbevolen om aan de puntmetingen zoals die nu aan het ISTM worden aangeboden eveneens geaggregeerde grootheden toe te voegen. Die geaggregeerde grootheden variëren van totale of gemiddelde transporten (grootte of richtingen) over een raai of een deelgebied, tot zelfs termen in balansvergelijkingen op het niveau van morfologische eenheden.

Bij de hierboven genoemde bronnen voor het inbrengen van meer fysische en/of data-georiënteerde voorkennis moet nog het volgende worden opgemerkt. In de praktijk zal het niet altijd mogelijk zijn om deze bronnen in een strikt kwantitatieve vorm voor te schrijven. Desalniettemin kunnen met geschikte ('educated') schattingen, of aan de hand van gerichte variaties de ISTM resultaten in een gewenste richting worden gestuurd. Op dat moment kunnen de hier genoemde voorbeelden van fysische of data georiënteerde voorkennis dan worden opgevat als fysische vormen van regularisatie. Dit laat zien dat het inbrengen van voorkennis en het toepassen van regularisatie niet geheel los van elkaar staan, maar geleidelijk in elkaar over kunnen gaan.

Regularisatie is een mechanisme om aan de hand van gladheidseisen toch een éénduidig transportenveld te kunnen afleiden wanneer de hoeveelheid fysische/meet-informatie daarvoor onvoldoende is. Zelfs indien conform het bovenstaande meer fysische kennis in het ISTM wordt ingebracht zal het naar verwachting nog steeds niet mogelijk zijn om zonder regularisaties het inverse probleem tot een goed einde te brengen. Aan de hand van een review van regularisatie in het algemeen, en een wiskundige analyse van de effecten van de tot nu toe gehanteerde vormen van regularisatie in het bijzonder, zijn een aantal aanbevelingen afgeleid voor het verbeteren van dit mechanisme. Een van de belangrijkste conclusies die hierbij is afgeleid is dat het hierin aanbrengen van ruimtelijke variaties belangrijke voordelen kan bieden t.o.v. het uniform toepassen zoals dat tot nu toe moest worden gedaan. Deze ruimtelijke variaties moeten worden gerelateerd met de verwachte, of aan de hand van de a priori waargenomen, spatiële variabiliteit van het transportenveld. Een tweede belangrijke verbetering kan worden bereikt door deze functionaliteit meer flexibel/generiek te maken in de zin dat de gebruiker zelf de vorm en de mate van regularisaties kan samenstellen in plaats van het moeten kiezen uit de nu strikt voorgeschreven vormen. Op dat moment kan regularisatie niet alleen verschillend in gewicht, maar tevens verschillend qua vorm per deelgebied, raai, traject of zelfs per individueel punt(en) worden opgelegd.

Voor het verbeteren van de wijze waarop in het ISTM informatie van *transportrichtingen* wordt verwerkt is het conceptuele model aangepast t.a.v. de richtingsvergelijkingen. Hierbij

zijn nu twee opties/varianten gedefinieerd. In ‘de strikt parallelle optie’ zal het ISTM streven naar het zoveel mogelijk exact doen samenvallen van een gemeten en de te schatten transportrichting. Een verschil van 180° van die hoeken (ofwel anti-parallelle richtingen) wordt dan gezien als de grootst mogelijke afwijking van het gewenste resultaat. In de alternatieve optie vervalt het onderscheid tussen parallel en anti-parallel. In dat geval wordt een verschil van 90° (ofwel onderling loodrechte richtingen) gezien als de grootst mogelijke afwijking van het gewenste resultaat. De eerste optie zal worden gebruikt als van transporten de exacte richting bekend is, terwijl de tweede optie gekozen zal worden als alleen de baan van een transport bekend is en daarmee de richting modulo 180° .

Een belangrijke konsekwentie van de voornoemde verbetering van het verwerken van transportrichtingen is dat het ISTM moest worden uitgebreid met niet-lineaire vergelijkingen. Op zijn beurt vergt deze uitbreiding een andere procedure voor het daarna oplossen van de ISTM-vergelijkingen. Deze aanpassingen voor het niet-lineair verwerken van richtingsinformatie zijn hier niet alleen uitgebreid beschreven maar ook in het ISTM geïmplementeerd, en aan de hand van een specifieke test-case getoetst. Dit heeft geleid tot de operationele versie 2.00 van het ISTM als opvolger van de vorige versie 1.03.

I Inleiding

In 1993-1994 is door WL | Delft Hydraulics in opdracht van Rijkswaterstaat/RIKZ het Invers Sediment Transport Model (ISTM) ontwikkeld voor het invers modelleren van twee dimensionale (horizontale) sedimenttransporten. Met dit lineaire ISTM kunnen diverse bronnen van informatie van sedimenttransporten worden geïntegreerd die voor een bepaald gebied beschikbaar zijn. Deze informatiebronnen zijn (i) gemeten data van transportvectoren, transportrichtingen, bergingen, en randvoorwaarden, (ii) fysische kennis in de vorm van de continuïteitsvergelijking waaraan de transporten en bergingen moeten voldoen, en (iii) wiskundige voorwaarden in de vorm van zogenaamde regularisatievergelijkingen waarmee de ruimtelijke variabiliteit van de transportvectoren kan worden gestuurd. Het resultaat van een ISTM-berekening is dan een transportenveld dat het over een tijdsinterval gemiddelde sedimenttransport weergeeft dat optimaal past bij de gegeven informatie. Dit transportenveld bestaat uit de transportvectoren

$$\vec{T}_m^n = \left((T_x)_m^n, (T_y)_m^n \right) \text{ voor alle cellen } (m,n) \text{ van een rechthoekig rooster waarmee het}$$

betreffende gebied spatieel gediskretiseerd is.

Conceptueel heeft het ISTM (versie 1.03) de vorm van een lineair stelsel vergelijkingen in de transportcomponenten $\left\{ (T_x)_m^n, (T_y)_m^n \right\}$ waarbij aan elke vergelijking een additieve ruis is

toegevoegd voor het weergeven van de onzekerheid (ofwel gewicht) van die vergelijking. De oplossing van dit stelsel lineaire stochastische vergelijkingen is gebaseerd op het "Maximum Likelihood" principe. Numeriek moet hiervoor een ander daarvan afgeleid lineair stelsel van vergelijkingen worden opgelost. Dit levert dan het gewenste transportenveld waarin op een statistisch consistente en optimale vorm alle informatiebronnen geïntegreerd zijn. Voor een meer gedetailleerde beschrijving van het onderliggende wiskundige en numerieke model, de oplosprocedure, en de resultaten van toepassingen wordt verwezen naar Koster en Van den Boogaard (1993ab), Van den Boogaard (1994), en Van den Boogaard en Mulder (1997). In de appendices van het rapport van Van den Boogaard (1994) is aan de hand van een beschrijving van de diverse invoer en uitvoer files een gebruikershand-leiding van het lineaire ISTM (versie 1.03) te vinden.

Het lineaire ISTM (versie 1.03) is door RIKZ inmiddels voor diverse praktijktoepassingen ingezet. Zie bijvoorbeeld Mulder (1996) voor zand- en slibbalansen voor het Eems-Dollard estuarium (uitgevoerd in het kader van het projekt DYNASTAR), en Van Marion (1999) voor zandbalansen voor het Zeegat van Texel (uitgevoerd in het kader van het K2000*KOP projekt). In deze grootschalige praktijktoepassingen werden zeer bevredigende resultaten gevonden. Desalniettemin geven de daarin opgedane ervaringen aanleiding om in nader detail na te gaan of het modelconcept van het ISTM op een aantal punten verbeterd zou kunnen worden. Expliciet wordt hierbij aan de volgende drie aspecten gedacht:

1. Het inbrengen van meer fysika in het model
2. De (her)definitie van de wijze waarop transportrichtingen in het model verwerkt worden
3. De functie van regularisatie in het ISTM

Voor een beter begrip van deze drie punten worden zij nu in nader detail toegelicht t.a.v. hun achtergrond, aanleiding en betekenis.

Ad 1.

Een met het ISTM gegenereerd transportenveld is in partiële mate gebaseerd op fysika. Het fysische fundament bestaat alleen uit de massabalansvergelijking, en dan alleen nog maar voor cellen of deelgebieden waarop gemeten bergingen beschikbaar zijn. Daarnaast kan fysische informatie van transportpatronen worden toegevoegd door het invoeren van gemeten groottes van transportcomponenten, of richtingen van het lokale transport. In de praktijk, zoals voor het hierboven genoemde Eems-Dollard estuarium en het Zeegat van Texel, zijn dergelijke puntmetingen meestal uiterst schaars. Dit geldt het meest uitgesproken voor transportgroottes. Hierdoor is onder praktische condities het totaal van de strikt fysische informatie onvoldoende om een éénduidig en consistent transportenveld voor het gehele domein af te leiden. Derhalve moet regularisatie worden toegepast om het inverse probleem alsnog op te kunnen lossen. In deze zin moet die regularisatie dus de informatie die ontbreekt in de fysische bronnen, en/of de onnauwkeurigheid van die informatie, compenseren. In de huidige vorm is deze regularisatie echter veel meer een mathematisch mechanisme dan dat deze een duidelijke fysische basis heeft. Alleen aan het feit dat regularisatie er voor zorgt dat de uiteindelijke oplossingen relatief continue en glad van vorm zijn mag een zekere, en in overmaat kwalitatieve, fysische betekenis aan dit mechanisme worden toegekend.

Het is duidelijk dat onder de hierboven genoemde condities van beperkte meetinformatie het uiteindelijk gevonden transportenveld in sterke mate bepaald kan zijn door de regularisatie. In het bijzonder geeft dit bij toepassingen het concrete probleem in welke mate en vorm regularisatie moet worden toegepast.

Deze achtergrond initieert dan de vraag of deze balans in fysische en mathematische informatie meer in de richting van de eerste kan worden verschoven door het inbrengen van meer fysika dan dat dit met de huidige ISTM-versie mogelijk is. Bij dergelijke kennis moet worden gedacht aan (a) *generieke proceskennis* van sedimenttransporten en/of morfologie zoals die is bevat in gangbare (empirische, hybride, conceptuele proces) modellen of relaties, of in de expertise van ervaren morfologen, of (b) *data-georiënteerde* kennis die specifiek ontleend wordt aan het domein waarop het ISTM wordt toegepast (zoals de geometrie, actuele meteo- en stromingscondities, effecten van ingrepen, etc.).

De zo ingebrachte fysische voorkennis wordt door het ISTM dan geïntegreerd met de data van de overige informatiebronnen. Bij de analyse/interpretatie van het dan met het ISTM gevonden transporten en bergingenveld moet dan wel worden gerealiseerd dat de daarin teruggevonden aspecten van de systeemkennis geheel of grotendeels het gevolg kunnen zijn van die toegevoegde ISTM-invoer. Worden op deze manier eigenschappen teruggevonden van *a priori onzekere* voorkennis dan mag dus aan de *a posteriori* resultaten dus geen waarheidsbewijs van die voorkennis worden ontleend. Anderzijds kan na de integratie van de fysische voorkennis en de overige (meet)data wel worden nagegaan óf, en in welke mate zij onderling consistent zijn. In dat geval is de 'waarheid' van die voorkennis nog steeds niet in strikte zin 'bewezen' maar kan misschien wel in voldoende mate aannemelijk worden gemaakt, om het toevoegen van die voorkennis aan de ISTM invoer te rechtvaardigen.

In meer concrete zin, en redenerend vanuit de praktijk en de specifieke doelstelling bij toepassingen, moeten bij het inbrengen van voorkennis in het bijzonder de tijd- en ruimteschalen en/of het wel of niet voldaan zijn aan evenwichtskondities worden betrokken bij de selectie en de te plegen voorbewerkingen van die voorkennis (alsmede de andere informatiebronnen) en het daarna kritisch interpreteren van de ISTM-resultaten. Deze kondities en hun consistentie bepalen in sterke mate de meerwaarde van de toegevoegde kennis.

Ad 2.

Bij het door het ISTM verwerken van gemeten transportrichtingen spelen in feite twee problemen. De eerste is dat binnen het oplossingsmechanisme er geen onderscheid gemaakt kan worden tussen een richting φ en de bijbehorende tegengestelde richting $\varphi + \pi$ (in radialen, of $\varphi + 180^\circ$ in graden). M.a.w., het ISTM is bij het vinden van de anti-parallelle richting net zo tevreden als bij het vinden van de gewenste richting. Een tweede, en minder ernstige, tekortkoming is dat de richtingsvergelijkingen niet ontkoppeld kunnen worden van de transportgroottes T_x , T_y . Het laatste betekent dat het ISTM afwijkingen van de gewenste richtingen het ergst vindt voor de grootste transporten. Van afwijkingen van de richtingen bij kleine transportgroottes trekt het model zich minder aan.

Er moet worden opgemerkt dat de oorspronkelijke wens om een lineair ISTM op te stellen debet is aan het voorafgaande: binnen de eis van lineariteit kan niet van de huidige geïmplementeerde richtingsvergelijkingen worden afgeweken, en kunnen als zodanig bovenstaande twee aspecten niet worden opgelost.

In de praktijktoepassingen is gebleken dat het ISTM op deze manier soms anti-parallelle richtingen genereert voor posities waar dat uitdrukkelijk niet gewenst is, of niet zou worden verwacht. Derhalve moet het huidige model voor het integreren van richtingsinformatie herzien worden om de dualiteit parallel/anti-parallel op te kunnen heffen. In feite is de gebruikerswens nog strikter van vorm gegeven het feit dat voor andere gevallen alleen de baan van transporten bekend is en de richting daarmee alsnog parallel of anti-parallel kan zijn. Als resultaat moeten voor het verwerken van richtingsinformatie dus twee criteria mogelijk zijn, namelijk strikt parallel enerzijds, en parallel óf anti-parallel anderzijds. Bovendien moeten de modelvergelijkingen voor die richtingscriteria uitgedrukt worden in hoeken zonder dat hierin (de a priori onbekende) transportgroottes in zijn betrokken. De consequentie van een uitbreiding van het ISTM dat beter aan deze criteria kan voldoen is dat dan wel dat op een niet-lineair model moet worden overgeschakeld. Dat heeft vervolgens weer repercussies op de te volgen oplosmethode.

Ad 3.

Zoals onder (1) al is vermeld is in de praktijk de hoeveelheid fysische/meet-informatie onvoldoende om met het ISTM éénduidige transportenvelden te genereren en moet derhalve regularisatie worden toegepast. Zelfs indien het mogelijk is om meer fysische kennis in het ISTM in te brengen, conform de lijn geschetst onder (1), zal het naar verwachting nog steeds niet mogelijk zijn om zonder regularisaties het inverse probleem tot een goed einde te brengen.

Op dit moment zijn in het ISTM 10 vormen van regularisatie op te geven. Deze bestaan uit eerste en tweede orde differenties van de transportcomponenten van naburige cellen. Deze differenties worden langs de x en y coördinaat-richtingen genomen en voor elke (m,n) -cel van het numerieke rooster toegepast. Bovendien wordt per vorm de regularisatie uniform toegepast, d.w.z. voor elke cel wordt dezelfde onzekerheid (ofwel gewicht) aan de betreffende vergelijking toegevoegd. Zie Koster en Van den Boogaard (1993ab) voor een meer gedetailleerde beschrijving van het regularisatiemechanisme.

De gewichten in de diverse regularisatievormen moeten aan de hand van gerichte, herhaalde ISTM-experimenten vastgesteld worden. Dit 'tunen' van die gewichten is in sterke mate een subjectief proces omdat aan de hand van het vergelijken van de subsessievelijk gevonden transportvelden de gebruiker een beslissing moet nemen welke dan uiteindelijk 'het beste' is.

Bij de nu gehanteerde vormen van regularisatie zijn diverse kanttekeningen/beperkingen te plaatsen. Op de eerste plaats kan het konstant zijn van de aan de regularisatievergelijkingen toegekende gewichten worden genoemd. Immers, gezien het smoothende effect van regularisatie mag men verwachten dat op deelgebieden met een grote variabiliteit van de

transporten minder regularisatie moet worden toegepast dan op deelgebieden waar het transportenveld veel meer uniform van vorm is. Deze suggestie voor niet uniforme regularisatie correspondeert met de bevindingen van Van Marion in zijn ISTM-toepassingen voor het Zeegat van Texel waarbij schattingen zijn gemaakt van die ruimtelijke variabiliteit van de transporten. Naast dit aspect kunnen in min of meer willekeurige volgorde als verdere kanttekeningen/ beperkingen worden genoemd dat het niet geheel duidelijk is wat de effecten (overeenkomsten/ verschillen) van de diverse regularisatievormen zijn, de ongevoeligheid van de diverse differenties voor transportgroottes, de subjectiviteit in het tunen van de gewichten, de oriëntatie van de differenties die strikt langs de x of de y coördinaatas gericht is en niet langs willekeurige richtingen kan worden gedefinieerd.

Dit alles leidt tot de suggestie om in nader detail te evalueren of met aanpassingen van de huidige vorm van regularisatie een verdere verbetering van het ISTM kan worden gerealiseerd. Er moet hierbij worden opgemerkt dat dit onderwerp niet geheel los kan worden gezien van de onder (1) beoogde doelstelling van het inbrengen van meer fysische kennis. Immers, ook bij het verbeteren van het regularisatiemechanisme bestaat een voorkeur voor vormen die, voor zover dat mogelijk is, zo expliciet mogelijk zijn afgestemd op (additionele) fysische-, proces-, of gebiedskennis. Op deze manier worden de oplossingen dan veel meer door fysika gestuurd dan door de nu min of meer willekeurige mathematische gladheidsvoorwaarden. In dat licht kunnen dan op een zeker moment vergelijkingen die uit (1) voortkomen eveneens als een regularisatie worden opgevat (of vice versa), vooral als deze vergelijkingen een hoge mate van empirie, aggregatie, en/of meer kwalitatieve dan kwantitatieve kennis representeren.

De drie hierboven genoemde en daarna qua achtergrond kort toegelichte aspecten voor verbetering van het ISTM worden in het vervolg van deze rapportage in nader detail uitgewerkt. Daartoe wordt in Hoofdstuk 2 aan de hand van een review van gangbare (empirische, hybride, en proces) sedimenttransport/morfologische-modellen geïnventariseerd welke fysische fundamenteën van deze modellen bruikbaar kunnen zijn om het bestaande ISTM met meer fysische kennis uit te breiden. Op dezelfde manier wordt in Hoofdstuk 3 nagegaan hoe voor zo'n uitbreiding eveneens data-georiënteerde voorkennis (d.w.z. case/gebieds-specifiek) zou kunnen worden gebruikt. Hoofdstuk 4 bevat een uitgebreide review van regularisatie waaraan vervolgens diverse ideeën worden ontleend voor het verbeteren van dit mechanisme in het ISTM.

Voor het verbeteren van de wijze waarmee richtingsinformatie in het ISTM wordt verwerkt moet het model naar een niet-lineaire vorm worden uitgebreid. Het conceptuele model van een generiek niet-lineair ISTM, en diens relatie met de huidige lineaire model, alsmede de procedure die moet worden gevolgd om de oplossing van het niet-lineaire model te vinden worden uitgebreid beschreven in Hoofdstuk 5. In het bijzonder worden dan in Hoofdstuk 6 de vergelijkingen gegeven waarmee in het niet-lineaire model informatie van transport-richtingen wordt weergegeven.

De voornaamste in dit onderzoek gevonden konklusies en aanbevelingen worden daarna kort samengevat in Hoofdstuk 7.

Van alle hier voorgestelde verbeteringen is op dit moment (Januari 2000) alleen de niet-lineaire uitbreiding t.a.v. het verwerken van gemeten richtingen daadwerkelijk in het ISTM geïmplementeerd. Deze implementatie is conform het model beschreven in Hoofdstuk 6, en is getoetst aan de hand van een expliciete case. Dit heeft geleid tot de release van een nieuwe (en niet lineaire) versie 2.00 van het ISTM. De hieruit voortkomende wijzigingen in het prepareren van invoerfiles, de feitelijke run procedure, en de resulterende wijzigingen in de uitvoerfiles worden kort beschreven in de vorm van een gebruikershandleiding die in

Appendix A van dit rapport te vinden is. In Appendix B zijn de resultaten van de test case beschreven waarmee het nieuwe richtingsformalisme getest is.

Naast de implementatie van een niet-lineair richtingsformalisme heeft in versie 2.00 van het ISTM eveneens een aanpassing van de diagnostische tool plaatsgevonden. Terwijl in versie 1.03 alle kostenfunctietermen werden berekend, en afgedrukt in de print file, wordt dat in versie 2.00 van het ISTM alleen nog gedaan voor de termen die door de gebruiker daadwerkelijk geactiveerd zijn. Voor een gedetailleerde beschrijving van de diagnostische tool en de daarvoor gedefinieerde kostenfuncties wordt verwezen naar Van den Boogaard (1994).

2 Fysische voorkennis

2.1 Empirische voorkennis

De in de inleiding beschreven fysische en morfologische systeemkennis kan worden onderverdeeld in:

- geometrische voorkennis, en
- morfologische evenwichtsrelaties.

I. De eerste categorie bevat de meest geaggregeerde informatie en dient direct door een ervaren morfoloog aan het ISTM aangeboden te worden. Bij deze *expert-kennis* wordt aan de volgende informatie gedacht:

- I.1. Het sedimenttransport in estuaria en zeekommen vindt hoofdzakelijk plaats in de geulen in hun langrichting en in de vorm van uitwisseling tussen geulen en intergetijdegebieden. Langtransport over de intergetijdegebieden is over het algemeen klein. Op deze wijze is de hoofdrichting van de transporten vast te leggen. Of de transporten im- of exporterend zijn is echter niet vast te stellen met het lineaire ISTM, versie 1.03 (transporten $\pm 180^\circ$).
- I.2. Deze informatie kan handmatig door een ervaren morfoloog aan het ISTM worden aangeboden. Een alternatieve aanpak is via een algoritme dat lokaal de grootste gradiënten in de bodemligging opzoekt; deze zijn representatief voor plaatgeulovergangen. Deze alternatieve aanpak vraagt om een robuust patroonherkenningsalgoritme en wordt daarom vooralsnog niet aanbevolen.
- I.3. Door bestudering van bathymetrische kaarten kan een ervaren morfoloog eb- en vloedgeulen en hun loop onderkennen. De kenmerkende karakteristieken van dergelijke geulen zijn indertijd al door van Veen (1950) onderkend. Het wordt aanbevolen ook deze kennis, die lokale informatie over de netto transportrichtingen levert, handmatig aan het ISTM aan te bieden.

II. De onder de *tweede categorie* vallende empirische kennis is onder meer samengevat door Allersma (1988) en Eysink (1990, 1992). Echter velen hebben zulke empirische relaties bestudeerd en er over gepubliceerd. Een overzicht van deze relaties werd o.a. door Wang et al. (1998) gegeven. Van al deze *systeemkennis* kunnen voor het ISTM de volgende relaties relevant zijn:

II.1 De afmetingen van de geuldoorsnede A_c in relatie tot de achterliggende komberging. Een bekende evenwichtsrelatie luidt:

$$A_c = c_1 P^n \quad (2.1)$$

waarin c_1 en n een constante ($n \approx 1$), A_c de geuldoorsnede onder gemiddeld water (MSL) en P het getijprisma ter plaatse van die doorsnede, i.e. het watervolume tussen hoog- en laag water. Merk op dat c_1 een niet-dimensieloze parameter is. Verder is van belang dat deze relatie geldt voor grotere, geaggregeerde gebieden.

Morfologische veranderingen, dus veranderingen in de berging van de kom impliceren niet automatisch variaties in het getijprisma. Omgekeerd zal een verandering in getijprisma altijd veroorzaakt worden door een toe- of afname in de berging, bij gelijkblijvende komgrootte. Indien bijvoorbeeld blijkt dat het getijprisma in de loop van de tijd afneemt, zal de komberging achter de voedende geul afnemen, en vice versa. Uit deze informatie kan echter geen informatie over de herkomst van het voor een verkleining van de berging benodigde sediment gedistilleerd worden. Wel kan deze informatie gebruikt worden om de aangeboden data op consistentie te toetsen (data validatie) en een betere onderbouwing van eventuele wegingsfuncties ter plaatse van die geul.

II.2 De hoogtes van de intergetijdegebieden, zoals de hoogte tussen gemiddeld laagwater (MLW) en MSL. Zulke relaties kunnen beschreven worden als:

$$\begin{aligned} H_{h,e} &= f_1(a_1, A_b) \\ H_{l,e} &= f_2(a_1, A_b) \end{aligned} \quad (2.2)$$

waarin $H_{h,e}$ en $H_{l,e}$ de evenwichtshoogte t.o.v. MSL weergeven van het intergetijdegebied onder, respectievelijk boven MSL, a_1 de getijamplitude en A_b de totale oppervlakte van het getijbassin. Deze informatie is slechts bruikbaar als ook de breedte van de intergetijdegebieden is gegeven (zie Eysink, 1992).

In tegenstelling tot Eysink geeft Siefert (1987a, 1987b) een continue beschrijving van de geometrie van intergetijdegebieden in de vorm van algebraïsche relaties. Deze evenwichtsbeschrijvingen zouden dus in principe geschikt zijn om aanvullende metingen voor het ISTM te genereren. Nadeel is echter dat deze relaties waarschijnlijk zeer gebiedsafhankelijk zijn (sedimentsamenstelling; golfklimaat, dus oriëntatie t.o.v. heersende windrichting; enz.). Daarom wordt aanbevolen om informatie over de vorm van intergetijdegebieden slechts te gebruiken voor data-validatie.

II.3 Op de schaal van het gehele bekken bestaan evenwichtsrelaties voor het totale areaal aan intergetijdegebieden A_{ig} :

$$A_{ig} / A_{MHW} = f_3(A_{MHW}) \quad (2.3)$$

waarin A_{MHW} het bassinoppervlakte onder gemiddeld hoogwater (MHW) is.

Aannemende dat deze relatie voor een systeem bekend is kan uit een vergelijking van de gemeten A_{ig} met haar evenwichtswaarde bepaald worden of het systeem te veel, dan wel te weinig sediment bergt. Hiermee kan dus de richting van het sedimenttransport in de voedende geulen worden bepaald.

Verder is het mogelijk om uit een vergelijking van de absolute waarde van de gemeten A en haar verandering over een zeker tijdsinterval met de evenwichtswaarde de meetdata te valideren voor gebruik in het ISTM, dan wel gewichtsfuncties aan de metingen bij gebruik in het ISTM toe te kennen.

Met behulp van deze relaties zou dus op tamelijk grote schaal (macro- tot mega-schaal) de richting van de sedimenttransporten bepaald kunnen worden. Voorzichtigheid is hierbij geboden: grote lokale morfologische verstoringen, bijvoorbeeld, worden dikwijls eerst

lokaal genivelleerd voordat het evenwicht wordt hersteld door sedimentimporten van, of -exporten naar buiten het systeem.

Sterke punten van het inbrengen van empirische kennis zijn:

- De geaggregeerde informatie is in eenvoudige vorm beschikbaar, en
- De informatie is overzichtelijk en goed toegankelijk.

Zwakke punten van het inbrengen van empirische kennis zijn:

- Empirische kennis levert geen informatie over transportgrootheden,
- Empirische kennis levert geen informatie over de relevante morfologische tijdschalen op, en
- Empirische kennis is dikwijls systeemafhankelijk en niet goed overdraagbaar naar andere systemen; ook is vaak niet bekend of een morfologisch systeem daadwerkelijk in evenwicht is.

2.2 Hybride modellen

Hybride modellen worden gekenmerkt door de integratie van empirische kennis van een systeem met deterministische, numerieke modellen voor het berekenen van de waterbeweging (en soms (aspecten van) het sedimenttransport). Deze modellen leveren typisch informatie over de morfologische tijdschalen en over de grootte en richting van het sedimenttransport en de grootte van bodemveranderingen op macro-schaal.

Ten behoeve van Rijkswaterstaat werden onder meer de volgende hybride modellen ontwikkeld:

ESTMORF (e.g. Wang et al., 1998) is een één-dimensionaal netwerkmodel opgezet voor het simuleren van (middel)lange-termijn morfologische ontwikkelingen in estuaria. Het model maakt gebruik van een één-dimensionaal waterbewegingsmodel en diverse empirische morfologische evenwichtsrelaties, zoals in Paragraaf 2.1 beschreven.

ASMITA (e.g. Stive et al., 1998) is een boxmodel opgezet voor het simuleren van lange-termijn morfologische ontwikkelingen voor estuaria, kustgebieden en voordelta's. De benodigde waterbewegingsinformatie wordt extern toegeleverd in de vorm van restdebieten en diverse empirische morfologische evenwichtsrelaties, vergelijkbaar aan de in ESTMORF gebruikte, worden toegepast.

In ISTM-terminologie zouden de volgende vergelijkingen aan het systeem worden toegevoegd:

$$\bar{T}_x + V_1 = \bar{\hat{T}}_x, \quad \bar{T}_y + V_2 = \bar{\hat{T}}_y, \quad \bar{T}_x \sin \bar{\hat{\varphi}} - \bar{T}_y \cos \bar{\hat{\varphi}} + V_3 = 0 \quad (2.4)$$

waarin de 'gedakte' grootheden de metingen, dus de parameters uit het hybride model vertegenwoordigen, V_i de toe te voegen ruis en de overgestreepte symbolen de geaggregeerde grootheden (i.e. geïntegreerd over een aantal ISTM-cellen). Merk op dat de drie ruistermen V_1 , V_2 en V_3 formeel niet onafhankelijk zijn, aangezien de transportgroottes en -richtingen sluitend zijn. Tenslotte kan worden opgemerkt dat de resultaten van hybride modellen als initiële meetdata behandeld kunnen worden.

Sterke punten van het inbrengen van kennis uit hybride modellen zijn:

- De empirische informatie is goed geordend beschikbaar.

- Bergingen en transporten vormen een consistent geheel; de massabalans is op elke schaal sluitend, en
- Met hybride modellen kan de morfologische evolutie worden voorspeld, zodat informatie van transporten en bergingen over elke tijdspanne beschikbaar is.

Zwakke punten van het inbrengen van kennis uit hybride modellen zijn:

- Hybride modellen zullen voor een nieuw systeem dikwijls opgezet en gecalibreerd dienen te worden.

2.3 Procesmodellen

Met het ISTM kunnen sedimenttransporten in de twee horizontale richtingen worden berekend. Daarom zal in deze paragraaf geen aandacht aan één-dimensionale sedimenttransportmodellen worden besteed.

De twee, twee-en-half- en drie-dimensionale procesmodellen kunnen onderverdeeld worden in twee categorieën:

1. Initiële modellen, dit zijn numerieke modellen waarmee, voor een gegeven waterbeweging en bodemligging, de sedimenttransporten en de initiële bodemveranderingen kunnen worden berekend en
2. Morfodynamische modellen, dit zijn numerieke modellen waarmee de bodemontwikkeling op lange termijn kan worden berekend, rekening houdend met de interactie tussen de actuele bodemligging, de waterbeweging en het sedimenttransport.

Het gebruik van twee- en twee-en-half-dimensionale initiële procesmodellen is min of meer state-of-the-art. Modellen vallend onder de tweede categorie worden ook wel quasi-3D modellen genoemd, omdat een aantal drie-dimensionale effecten geparameteriseerd worden meegenomen. Dit zijn (zie bijvoorbeeld WL | Delft Hydraulics, 1996 en Roelvink en Wang, 1996):

- Relaxatie-effecten voor zwevend sediment (de zgn. Galappatti-methode),
- De effecten van bochtstroming op de waterbeweging (secundaire stromingen),
- De effecten van bodemhellingen op het sedimenttransport, en
- Golf-stroominteractie in de kustzone.

Merk op dat met volledige (d.i. nog steeds hydrostatisch) drie-dimensionale procesmodellen nog weinig ervaring is opgedaan (e.g. Thoolen en Wang, 1999).

Het gebruik van procesmodellen impliceert dat simulaties met zulke modellen (bijvoorbeeld DELFT2D-MOR) uitgevoerd dienen te worden voor condities die maatgevend zijn voor de tijdsperiode waarover bodemveranderingen (de bergingen in het ISTM) beschikbaar zijn (zie bijv. Latteux, 1995). Hierbij dient gedacht te worden aan springtij-doodtijcycli of zogenaamde morfologische getijden. Hoe slechter dergelijke condities bekend zijn, hoe groter de variatie in de ruistermen van de continuïteitsvergelijking in het ISTM.

Met een sedimenttransportmodel kan per roosterpunt van een geschematiseerde gebied het netto sedimenttransport (grootte en richting) en de netto bodemverandering berekend worden. Na interpolatie naar het ISTM-rooster kunnen deze gegevens als meetwaarden aan het ISTM worden toegeleverd.

Merk op dat, zeker in geval van bodemtransport, het lokale sedimenttransport erg gevoelig is voor de lokaal berekende stroomsnelheid, dus voor de lokale initiële bodemligging (zie bijv. Thoolen en Wang, 1999). Dit kan tamelijk grote, kortgolfige fouten in de berekende transportvelden genereren (om deze reden zou het gebruik van een morfodynamisch model

te prefereren zijn, omdat in zo'n model initiële bodemonregelmaticheden worden weggefilterd). De fouten in de horizontale transportgradiënten kunnen zelfs nog groter zijn, afhankelijk van de lokale condities, en gebruik van gegevens van procesmodellen in de regularisatietermen (eerste en tweede orde afgeleiden) dient daarom ontraden te worden.

Aanbevolen wordt om eerst goed te verifiëren of de met een initieel procesmodel berekende transporten en bergingen de goede tendenzen weergeven en de data vervolgens te filteren en/of te integreren, alvorens deze aan het ISTM als "metingen" aan te bieden. Op deze wijze kunnen transportdata en sedimentbudgetten over morfologische eenheden worden verkregen welke aan het ISTM kunnen worden aangeboden (Jeuken, 2000).

In ISTM-terminologie zouden de volgende vergelijkingen aan het systeem worden toegevoegd:

$$T_x + V_1 = \hat{T}_x, \quad T_y + V_2 = \hat{T}_y, \quad T_x \sin \hat{\phi} - T_y \cos \hat{\phi} + V_3 = 0 \quad (2.5)$$

waarin de gedakte grootheden de metingen, dus de parameters uit het transportmodel vertegenwoordigen en V_i de toe te voegen ruis.

Sterke punten van het inbrengen van kennis uit initiële procesmodellen zijn:

- De modellen zijn min of meer operationeel en state-of-the-art,
- Bergingen en transporten vormen een consistent geheel; de massabalans is op elke schaal sluitend, en
- De informatie is op elke schaal beschikbaar.

Zwakke punten van het inbrengen van kennis uit initiële procesmodellen zijn:

- De initiële bodemligging past, ten gevolge van onnauwkeurigheden, vaak niet bij de berekende sedimenttransportvelden, met als gevolg onnatuurlijke transportgradiënten (en dus bodemveranderingen),
- De waterbeweging zelf, met name de voor resttransport zo belangrijke hogere harmonische componenten, is vaak niet voldoende nauwkeurig bekend,
- Ondanks veel onderzoek zijn de gangbare sedimenttransportformules nog steeds niet erg nauwkeurig.

Morfodynamische modellen bevatten gekoppelde elementen, waarbij de cyclus bodemligging - waterbeweging - sedimenttransport - nieuwe bodemligging enige (vele) keren geheel of gedeeltelijk wordt doorlopen, vaak tot een (nieuw) evenwicht is bereikt. De wijze waarop deze cyclus wordt doorlopen hangt af van de voor het betreffende systeem relevante ruimte- en tijdschalen (De Vriend et al., 1993). Morfodynamische modellen zijn volop in ontwikkeling en morfodynamisch modelleren is beslist nog geen state-of-the-art (De Vriend, 1997, 1998), hoewel tegenwoordig veelbelovende resultaten worden verkregen. In een studie met DELFT2D-MOR bijvoorbeeld voor het Amelandse Zeegat (Alkyon en WL | Delft Hydraulics, 1999, Roelvink, 1999) werd de migratie van geulen en de ontwikkeling van een buitendelta kwalitatief goed voorspeld. De kwantitatieve resultaten, zowel grootte als richting bleken echter gevoelig voor de gekozen hydro-meteo condities waarvoor de simulaties zijn uitgevoerd, en voor de sedimenttransportformuleringen. Daardoor zijn zij (nog) niet geschikt om als meetdata aan het ISTM toe te leveren.

Sterke punten van het inbrengen van kennis uit morfodynamische modellen zijn:

- bergingen en transporten vormen een consistent geheel; de massabalans is op elke schaal sluitend,
- de informatie is op elke ruimte- en tijdschaal beschikbaar, en
- de resultaten van morfodynamische modellen hebben geen hinder van lokale onregelmaticheden in de initiële bodemligging, mits de inspeelperiode lang genoeg is.

Zwakke punten van het inbrengen van kennis uit morfodynamische modellen zijn:

- morfodynamische modellen zullen voor elk systeem dikwijls opnieuw opgezet en gecalibreerd dienen te worden; dit betekent over het algemeen een forse inspanning, en
- morfodynamisch modelleren is nog geen ‘state-of-the-art’.

De zwakke punten wegen in dit geval zwaarder dan de sterke, en het gebruik van morfodynamische modellen om metingen voor het ISTM te genereren, wordt daarom vooralsnog afgeraden. Bij gebruik van data van een initieel model dient dit model wel in redelijke mate getoetst te worden.

2.4 Aanbevelingen

Voorgesteld wordt om, in volgorde van prioriteit, de volgende informatie als systeemkennis of extra metingen aan het ISTM toe te voegen.

1. De in Paragraaf 2.1 genoemde expert-kennis van een systeem,
2. De onder II.3 in Paragraaf 2.1 genoemde systeemkennis

De bij deze twee voorafgaande punten additionele informatie bestaat daarbij uit transport-richtingen. De bijbehorende vergelijkingen zijn voor het (niet-lineaire) ISTM zoals beschreven in Hoofdstuk 6 van dit rapport.

3. De in Paragraaf 2.3 genoemde transporten en bergingen verkregen met initiële procesmodellen, en
4. De in Paragraaf 2.2 genoemde transporten en bergingen verkregen met hybride modellen.

In verdere studies en toepassingen dient te worden uitgezocht welke van deze aanbevelingen de beste perspectieven bieden voor een daadwerkelijke verbetering van het ISTM.

3 Data-georiënteerde voorkennis

In Hoofdstuk 2 is nagegaan hoe in het ISTM additionele informatie kan worden ingebracht die aan proceskennis is ontleend. Die proceskennis kan variëren van gedetailleerde, generieke, fysische kennis (en bijvoorbeeld gekondenseerd in een procesmodel) tot meer geparametriseerde beschrijvingen van de fysika in de vorm van empirische relaties of zelfs vuistregels. Terwijl daar dus vooral van de fysika wordt uitgegaan, en daarmee meer generieke eigenschappen van sedimenttransporten of morfologie omvat, wordt in dit hoofdstuk meer vanuit de data en gebiedskennis uitgegaan en is daarmee specifiek voor de feitelijke case waar het ISTM moet worden toegepast.

Uiteraard is het ISTM zoals dat nu wordt toegepast al in sterke mate op data gebaseerd. Die data worden echter nu vooral als puntbronnen aan het model aangeboden. Dat komt er op neer dat voor elke (individuele) meting van een berging, transportgrootte of transportrichting een modelvergelijking wordt opgesteld die geassocieerd is met één bepaalde cel van het gebruikte rooster. Er is geen enkele reden om deze aanpak te verlaten of om puntmetingen in de toekomst op een andere manier dan nu te gaan verwerken. Het is meer de bedoeling om na te gaan in hoeverre het perspectieven kan bieden om aan het ISTM eveneens (ruimtelijk) *geaggregeerde* informatie, voorkennis, voorwaarden etc. toe te voegen. Geaggregeerd betekent dan dat de informatie met diens vertaling naar de bijbehorende modelvergelijking betrekking kan hebben op combinaties van de transporten en/of bergingen van een raai, een deelgebied, of zelfs het gehele modelgebied. Dergelijke combinaties zijn dan spatieel gedistribueerde grootheden.

Hieronder worden t.a.v. zo'n aggregatie een aantal eerste ideeën geïnventariseerd. Deze zijn zoals in de vorige alinea al is aangegeven in sterke mate ontleend aan een of andere selectie of partitie van het ISTM-domein in karakteristieke deelgebieden (of hun randen, of andere trajectoria) waarvoor een of andere 'geïntegreerde', 'kumulatieve' of 'gemiddelde' uitdrukking in transporten en/of bergingen wordt opgesteld. Door deze uitdrukkingen (met een bepaalde onzekerheid) in het ISTM met de overige (punt)informatie te integreren kan het af te leiden transportenveld dus meer globaal of gebiedsgewijze door de gebruiker worden gestuurd. Bij deze 'definitie' en 'doelstelling' van aggregatie kunnen dan de volgende drie opmerkingen worden geplaatst:

1. Zo'n aanpak, met in het bijzonder de selectie van deelgebieden, wordt sterk bepaald door de specifieke case, en de daarvoor beschikbare gebiedskennis, zoals de geometrie en stromingscondities. Derhalve kan aggregatie in de hier beschreven vorm zeker als het inbrengen van *data georiënteerde voorkennis* worden opgevat.
2. In de huidige implementatie van het ISTM is het nog niet mogelijk om geaggregeerde informatie aan te bieden en te verwerken. De hiervoor benodigde uitbreiding lijkt echter relatief eenvoudig realiseerbaar. Schetsmatig moet dat als volgt. Elke geaggregeerde grootheid (zoals de voorbeelden hieronder gegeven) kan wiskundig worden vertaald naar een vergelijking van de vorm: $F_\ell(\Theta_{k_1}, \Theta_{k_2}, \dots, \Theta_{k_N}) + V_\ell = 0$. Hierin zijn de Θ_{k_n}

componenten van de transport vector $\vec{\Theta}$, en V_ℓ de onzekerheid van de ℓ^e vergelijking met spreiding σ_ℓ . Nieuw t.o.v. het huidige conceptuele ISTM model is dan dat de functie $F_\ell(\cdot)$ *niet-lineair* kan zijn, en dat diens argumenten Θ_{k_n} veel verder van elkaar verspreid kunnen liggen dan naburige cellen. Een voorbeeld waarin men met een niet-lineaire vorm $F_\ell(\cdot)$ van doen heeft is het geval dat men richtingsinformatie (per cel, of geaggregeerd over meerdere cellen) in één vergelijking verwerkt.

Voor het feitelijk uitbreiden van het ISTM naar een niet-lineaire vorm wordt dan de oplossingsprocedure vertaald naar een optimalisatieprobleem. Dat wordt uitvoerig toegelicht in Hoofdstuk 5.

3. Het inbrengen van geaggregeerde grootheden kan worden gezien als een vorm van regularisatie ontleend aan gebiedskennis. Dit zal in Hoofdstuk 4 in de algemene review van regularisatie nader aan de orde komen.

Op dit moment kunnen dan puntgewijs de volgende eerste ideeën voor het inbrengen van geaggregeerde, en op data of gebiedskennis georiënteerde informatie worden genoemd.

- i. Op basis van de geometrie (geulen en platen systemen, getijbassins, gesteldheid van de bodem) en kennis van stromings- en/of meteo condities kan men voor individuele punten of gebieden mogelijk (verwachte) richtingen van transporten afleiden, of anderzijds gebieden aanwijzen waar transporten klein zullen zijn, of juist heel groot.
- ii. Een zeer bruikbare vorm van aggregatie kan bestaan uit het opstellen van een meetvergelijking die het totale transport over een raai geeft (of equivalent: het gemiddelde transport over zo'n raai). Zo'n raai zou bijvoorbeeld een open rand kunnen zijn waarover men de totale import of export voorschrijft.
- iii. Analooq aan (ii) kunnen voor sommige raaien misschien niet zozeer transportgrootten dan wel (voorkeurs)richtingen worden voorgeschreven. Hierbij wordt b.v. aan geulen gedacht.
- iv. Een directe generalisatie van (ii,iii) is het voorschrijven van een gemiddelde transport of richting over een deelgebied.
- v. Een raai zoals onder (ii) bedoeld kan een kromme en/of gesloten traject vormen die de rand vormt van een bepaald deelgebied. Aan de hand van een kringintegraal kan een vergelijking voor de totale berging worden opgesteld. Uit de gemeten bergingen binnen dat traject kan een 'rechterlid' worden geschat. Op deze manier kan men voor één of meerdere gebieden een continuïteitsvergelijking in geaggregeerde vorm worden opgesteld, naast de lokale vormen zoals die nu al voor elke cel aanwezig zijn. Die geaggregeerde vorm van de berging zal in het algemeen bestaan uit de som of gemiddelde van een groot aantal gemeten (punt)bergingen. Aangezien optellen, resp. middelen, onzekerheden reduceert is de geaggregeerde vorm naar verwachting representatiever en consistent met de werkelijkheid. Afhankelijk van de betrouwbaarheid van de beschikbare bergingen zou men dan zelfs kunnen overwegen om weinig of geen lokale vormen van de continuïteitsvergelijking meer te nemen. In plaats daarvan moeten dan een voldoende groot aantal over geschikte deelgebieden geaggregeerde vormen worden verwerkt. Op deze manier gebruikt men dan feitelijk gesmoothde vormen van de bergingen.
- vi. Het bovenstaande kan ook worden geïnterpreteerd in de vorm dat ook op de schaal van het hele systeem aan (bijvoorbeeld) de continuïteitsvergelijking dient te worden voldaan: het resulterend transport over de randen van dat systeem dient gelijk te zijn aan de verandering in de berging van het totale systeem. Met deze kennis kunnen bovengrenzen aan bijvoorbeeld het resulterend transport over de systeemranden gesteld worden. In ISTM-terminologie:

$$\oint_{A_b} \vec{T} \cdot \vec{n} \, ds = \iint (\hat{B} + \Delta B) \, dA$$

Hierin is $\vec{n}$ de normaalvektor op de systeemrand s en ΔB de bekende of geschatte overall fout in de gemeten berging.

- vii. Een speciaal geval van de voorafgaande aandachtspunten is dat men voor een raai of een deelgebied niet zozeer het totale of gemiddelde transport opgeeft, maar slechts de eis of dat gebied netto importerend of netto exporterend is.
- viii. Voor een deelgebied zou men kunnen voorschrijven welke fractie van de totale export van sediment in welk ander naburig deelgebied terecht moet komen. Een zelfde sturing kan voor importen worden opgesteld.

Bij deze inventarisatie wordt het volgende nog opgemerkt. Hoewel zeer gewenst, zal het in de praktijk niet altijd gemakkelijk zijn om op deze manier *kwantitatieve* transportinformatie te vergaren uit data-georiënteerde kennis en/of gebiedskennis. Vaak zal men niet verder komen dan indicaties of ruwe schattingen. Desalniettemin heeft het dan zin om dergelijke informatie aan ISTM-berekeningen toe te voegen. Het idee is dat aan de hand van herhaalde ISTM-berekeningen met daarin geschikte variaties van de aan gebieds/data-georiënteerde voorkennis ontleende (al dan niet geaggregeerde) invoergrootheden, veel inzicht kan worden verkregen in de lokale en globale gevoeligheid van de ISTM-resultaten voor die variaties. Met een kenners-oog kan dan bovendien uit de diverse zo gegenereerde transportvelden een meest realistische variant geselecteerd worden. Op deze manier heeft men dan op 'inverse' manier tevens een keuze gemaakt voor de invoerparameters waarvoor die variaties werden uitgevoerd.

4 Regularisatie-vergelijkingen

In de huidige versie van het ISTM, en voor de condities waarvoor het tot nu toe in de praktijk is ingezet, is zonder regularisatie het aantal (meet)vergelijkingen veel kleiner dan het aantal onbekenden, en is er geen éénduidige oplossing af te leiden. Met regularisaties worden dan additioneel gladheidseisen opgelegd waarna het oorspronkelijke slecht-gestelde inverse probleem wel (of meer) éénduidig kan worden opgelost.

Regularisatie wordt dus gebruikt om te weinig meetinformatie en/of modelvergelijkingen te compenseren. In de praktijk is het echter een lastig probleem in welke vorm en welke grootte regularisatie moeten worden toegepast. In het huidige 2D ISTM zijn op dit moment tien vormen gedefinieerd en deze zijn uniform t.a.v. de toe te voegen onzekerheden. Deze aspecten van deze vormen in het bijzonder, alsmede regularisatie in een brede algemene zin, worden in dit hoofdstuk nader onder de loep genomen. Daartoe wordt eerst in Paragraaf 4.1 kort samengevat hoe regularisatie in algemene zin wordt toegepast. Voor het ISTM wordt dit mechanisme in meer kwantitatief detail herbeschouwd, namelijk hoe vanuit een systeemtheoretisch oogpunt de tot nu toe gebruikte regularisatie als filter/interpolator van de gegeven metingen funktioneert. Uit deze *spektrale analyse* kan o.a. tevens de zin van niet uniforme regularisatie worden afgeleid, gegeven de (al dan niet apriori verwachte) spatiële variabiliteit van het transportenveld waaronder scherpe overgangen. Daarnaast lijkt het op deze manier mogelijk om in meer kwantitatieve zin voorschriften af te leiden voor de spreidingen die in de regularisaties moeten worden gebruikt om een bepaalde smoothing te realiseren. Additioneel aan de ideeën die ontleend zijn aan de in de Paragrafen 4.1 en 4.2 gepresenteerde meer algemene review van regularisatie, worden in Paragraaf 4.3 een aantal andere specifieke suggesties uitgewerkt die voor het ISTM perspectieven lijken te bieden t.a.v. smoothing. In Paragraaf 4.4 worden dan de voor de ISTM-praktijk meest relevant geachte vormen van regularisatie kort en puntsgewijs samengevat.

4.1 Review achtergrond van regularisatie

In deze paragraaf wordt een korte review gegeven van het toepassen van regularisatie in inverse problemen. Bij het laatste wordt meteen gekozen voor de vorm waarin een kostenfunctie $\tilde{\Theta} \rightarrow J(\tilde{\Theta})$ moet worden geminimaliseerd. Wanneer dit een niet goed gesteld probleem vormt (meervoudige of zelfs een oneindig aantal oplossingen) kan het gewenst zijn (zowel vanuit conceptueel als vanuit numeriek oogpunt) om additionele informatie (criteria, vergelijkingen, constraints, etc.) toe te voegen om het probleem wel éénduidig te kunnen oplossen. Dat komt er dan op neer dat van alle oplossingen degene die het beste voldoet aan de additionele eisen de voorkeur krijgt.

De regularisatie bestaat er dan uit om de additionele eisen weer te geven met, zeg, N vergelijkingen in de te schatten grootte $\tilde{\Theta}$. Deze vergelijkingen worden geacht te zijn herleid tot een vorm met 0 in het rechterlid. Voor het kunnen omgaan met onzekerheden in - (ofwel wegingen van) deze vergelijkingen wordt aan elke vergelijking een ruis met gemiddelde 0 toegevoegd. De n^e vergelijking ziet er dan als volgt uit:

$$H_n(\tilde{\Theta}) + V_n = 0 \quad (4.1.1)$$

In vektorvorm luidt het stelsel regularisatievergelijkingen dan:

$$\vec{H}(\vec{\Theta}) + \vec{V} = 0 \quad (4.1.2)$$

Het voorafgaande kan nu worden samengevat met de opmerking dat regularisatie in feite een vorm van (additioneel) modelleren is. Dit modelleren heeft betrekking op de informatie die men wil toevoegen aan de in principe fysisch georiënteerde informatie bevat in de oorspronkelijke kostenfunctie $J(\cdot)$. Een belangrijk punt is nu het op een *consistente* manier toevoegen van de regularisatievergelijkingen van (4.1.1) aan die oorspronkelijke kostenfunctie. Refererend aan het Maximum Likelihood formalisme betekent consistent dat bij het toevoegen van een regularisatievergelijking een gewicht aan die vergelijking moet worden toegekend dat afgeleid is van de onzekerheid V_n van die vergelijking. Dat is echter nog niet alles. Indien de onzekerheden V_n van twee of meer van de in totaal N vergelijkingen afhankelijk zijn dan moet de weging van die vergelijkingen daarmee overeenkomstig zijn, en moeten die vergelijkingen in gekombineerde vorm aan de oorspronkelijke kostenfunctie $J(\cdot)$ worden toegevoegd. Zonder nu verder op de mathematische details in te gaan kan dan worden aangetoond dat de korrekte uitbreiding van de kostenfunctie $J(\cdot)$ naar een kostenfunctie $L(\cdot)$ met daaraan de regularisatievergelijkingen met de juiste combinatie van gewichten toegevoegd er als volgt uit ziet:

$$L(\vec{\Theta}) = J(\vec{\Theta}) + \vec{H}^T(\vec{\Theta}) \cdot R^{-1} \cdot \vec{H}(\vec{\Theta}) = J(\vec{\Theta}) + \sum_{\mu=1}^N \sum_{\nu=1}^N H_{\mu}(\vec{\Theta}) (R^{-1})_{\mu\nu} H_{\nu}(\vec{\Theta}) \quad (4.1.3)$$

Hierin is R^{-1} de inverse van de covariantiematrix R van de ruizen. De covariantiematrix R is van dimensie $N \times N$ en is gedefinieerd door:

$$R := E[\vec{V} \cdot \vec{V}^T] \quad (4.1.4a)$$

In scalar-vorm betekent dit dat het matrix element $R_{\mu\nu}$ ‘bij de entry (μ, ν) ’ ($1 \leq \mu \leq N$, $1 \leq \nu \leq N$) is gegeven door:

$$R_{\mu\nu} := E[V_{\mu} V_{\nu}] \quad (4.1.4b)$$

Het symbool ‘:=’ in de vergelijkingen (4.1.4ab) betekent dat het symbool aan de linkerzijde is gedefinieerd met de expressie aan de rechterzijde.

De moraal is dus dat bij de kwadratische uitbreiding van $J(\cdot)$ niet zonder meer het kwadraat van de linkerleden van de regularisatievergelijkingen moet worden genomen en dan bij $J(\cdot)$ opgeteld, maar dat al deze vergelijkingen moeten worden gewogen conform de covariantiestructuur (=afhankelijksstructuur) van de bijbehorende onzekerheden. Het standaard kwadratisch optellen is dus alleen toegestaan als de diverse onzekerheden V_n onafhankelijk (formeel: ongecorreleerd) van elkaar zijn. Immers in dat geval zijn alleen de diagonaal elementen van R ongelijk aan nul, en zijn deze van de vorm: $R_{\mu\mu} =: \sigma_{\mu}^2$. Het matrix-element $R_{\mu\nu}$ kan dan worden geschreven als $R_{\mu\nu} = \delta_{\mu\nu} \cdot \sigma_{\mu}^2$ waarbij de Kronecker

$\delta_{\mu\nu} = 0$ voor $\mu \neq \nu$, en $\delta_{\mu\nu} = 1$ voor $\mu = \nu$. In dat geval heeft vanwege de diagonaal structuur van R de inverse ook een diagonaal structuur volgens: $(R^{-1})_{\mu\nu} = \delta_{\mu\nu} \cdot \frac{1}{\sigma_\mu^2}$. Invullen van deze vorm in vergelijking (4.1.3) geeft dan:

$$L(\Theta) = J(\Theta) + \sum_{\mu=1}^N \sum_{\nu=1}^N H_\mu(\bar{\Theta}) \cdot \left(\frac{1}{\sigma_\mu^2} \cdot \delta_{\mu\nu} \right) \cdot H_\nu(\bar{\Theta}) = \sum_{\mu=1}^N \left(\frac{H_\mu(\bar{\Theta})}{\sigma_\mu} \right)^2 \quad (4.1.5)$$

Inderdaad wordt dan een standaard kleinste-kwadratenform gevonden. Om de consistentie van de vergelijkingen (4.1.3) (voor afhankelijke ruizen) en (4.1.5) (voor onafhankelijke ruizen) te benadrukken wordt het volgende opgemerkt. De stochastisch afhankelijke modelvergelijkingen (4.1.2) kunnen naar een equivalent stelsel met stochastisch onafhankelijke vergelijkingen getransformeerd worden door linker en rechterlid beide met de matrix $R^{-\frac{1}{2}}$ te vermenigvuldigen. Dat geeft dan:

$$R^{-\frac{1}{2}} \cdot \bar{H}(\bar{\Theta}) + R^{-\frac{1}{2}} \cdot \bar{V} = 0 \quad (4.1.6a)$$

Ofwel:

$$\tilde{\bar{H}}(\bar{\Theta}) + \tilde{\bar{V}} = 0 \quad (4.1.6b)$$

Dat de getransformeerde vergelijkingen (4.1.6b) stochastisch onafhankelijk zijn volgt uit het feit dat de covariantiematrix van de getransformeerde ruizen $\tilde{\bar{V}} := R^{-\frac{1}{2}} \cdot \bar{V}$ de éénheidsmatrix $I = \{\delta_{\mu\nu}\}_{\mu,\nu=1}^N$ oplevert. Immers:

$$E[\tilde{\bar{V}} \cdot \tilde{\bar{V}}^T] = E\left[R^{-\frac{1}{2}} \cdot \bar{V} \cdot \left(R^{-\frac{1}{2}} \cdot \bar{V}\right)^T\right] = R^{-\frac{1}{2}} \cdot E[\bar{V} \cdot \bar{V}^T] \cdot R^{-\frac{1}{2}} = R^{-\frac{1}{2}} \cdot R \cdot R^{-\frac{1}{2}} = I \quad (4.1.7)$$

Uitgaande van de stochastisch onafhankelijke vergelijkingen (4.1.6b), en met de éénheidsmatrix voor de covariantiematrix van de ruizen $\tilde{\bar{V}}$, kan dan de standaard kwadratische uitbreiding van $J(\cdot)$ worden uitgevoerd zoals met vergelijking (4.1.5) is aangegeven om te vinden dat:

$$\begin{aligned} L(\bar{\Theta}) &= J(\bar{\Theta}) + \tilde{\bar{H}}^T(\bar{\Theta}) \cdot \tilde{\bar{H}}(\bar{\Theta}) = \\ &= J(\bar{\Theta}) + \left(R^{-\frac{1}{2}} \cdot \bar{H}(\bar{\Theta})\right)^T \cdot \left(R^{-\frac{1}{2}} \cdot \bar{H}(\bar{\Theta})\right) = \\ &= J(\bar{\Theta}) + \bar{H}^T(\bar{\Theta}) \cdot R^{-1} \cdot \bar{H}(\bar{\Theta}) \end{aligned} \quad (4.1.8)$$

Op deze manier wordt dan uiteindelijk toch weer de $L(\cdot)$ van vergelijking (4.1.3) terug gevonden zoals geponeerd voor het algemene geval van afhankelijke ruizen.

Tot zover de formele achtergrond van het op een consistente manier integreren (ofwel uitbreiden) van een kostenfunctie met regularisatievergelijkingen. We keren nu terug naar de ISTM-praktijk en evalueren het voorafgaande t.o.v. de tot nu toe gehanteerde vormen van regularisatie. Daarna kunnen we (eventueel) de stap zetten naar meer algemene/generieke vormen van regularisatie.

Voor het evalueren en adequaat kunnen interpreteren van de tot nu toe gehanteerde regularisaties beschouwen we vooralsnog een *ID* aanpak en concentreren ons op de eerste orde regularisaties zoals die tot nu toe in het lineaire ISTM (versie 1.03) zijn geïmplementeerd. Voor transporten T_k gedefinieerd op een rooster x_k ($1 \leq k \leq K$) luidt de bijdrage van de eerste orde regularisatie aan de kostenfunctie:

$$L_1 = \sum_{k=2}^K \left(\frac{T_k - T_{k-1}}{\Delta x \cdot \sigma_1} \right)^2 \quad (4.1.8)$$

Aangezien er dus voor elke $\ell \neq k$ kruistermen van de vorm $(T_\ell - T_{\ell-1}) \cdot (T_k - T_{k-1})$ ontbreken in de L_1 van vergelijking (4.1.8) betekent dit dat verondersteld is dat de onzekerheden in de differenties $T_\ell - T_{\ell-1}$ en $T_k - T_{k-1}$ onafhankelijk van elkaar zijn. Dat is zelfs voor $\ell = k + 1$ het geval, ook al komt T_k dan in beide uitdrukkingen voor. Een nadere uitgebreide analyse waarvan we de details hier weglaten laat dan zien dat dit model voor de onzekerheden van de diverse differenties $T_k - T_{k-1}$ kan worden vertaald naar het volgende model voor de onzekerheden V_k in de transporten T_k ‘an sich’:

$$\forall_{1 \leq k \leq K}: T_k = T_0 + V_k \quad , \quad \forall_{1 \leq k, \ell \leq K}: E[V_k V_\ell] = \sigma_0^2 \quad (4.1.9)$$

Het predicaat ‘ $\forall_{1 \leq k \leq K}$ ’ moet hierin worden gelezen als ‘voor alle k met $1 \leq k \leq K$ ’. In woorden betekent vergelijking (4.1.9) dat het aan de eerste orde regularisatie ten grondslag liggende model veronderstelt dat alle transporten eenzelfde en vast (maar onbekend) gemiddelde T_0 hebben, en waarbij bovendien alle onzekerheden maximaal wederzijds gekorreleerd zijn, d.w.z. alle ruizen V_k zijn gelijk. Hoewel dit (op dit moment, en dus achteraf gezien) geen inconsistentie omvat is dit foutenmodel erg rigide, temeer wanneer dit wordt toegepast op het gehele modelgebied. Derhalve kan de huidige vorm van eerste orde regularisatie weliswaar gehandhaafd blijven, maar lijkt het beter om deze op deelgebieden toe te passen waarop konstantheid (of beter gezegd beperkte variabiliteit) van de transporten en hun onzekerheden een redelijke aanname is. Op verschillende deelgebieden kan dan een afzonderlijke spreiding σ_1 worden toegekend, terwijl zij ontkoppeld kunnen worden door aan de randen of de complementen van die gebieden een relatief grote spreiding te kiezen, of de regularisaties aldaar in hun geheel niet in de uitgebreide kostenfunctie op te nemen. Al met al pleit dit dus voor plaatsafhankelijke regularisaties. Bij dit alles moet wel worden aangetekend dat het in de praktijk vaak moeilijk zal zijn om een ‘beperkte variabiliteit’ en de daarvoor dan te kiezen spreidingen te kwantificeren. Net zoals nu dienen dan weer meerdere experimenten met variaties van die spreidingen te worden uitgevoerd.

Dezelfde overwegingen die hier zijn gegeven voor de eerste orde regularisatie kunnen eveneens worden afgeleid voor de tweede orde regularisaties.

Zoals in het voorafgaande betoog al is gekonstateerd zijn de tot nu toe gehanteerde eerste en tweede orde regularisaties *speciale* vormen, die bovendien telkens worden toegepast op basis van *naburige cellen*. Het is dan de vraag wat dan een meer algemene basis van (nulde, eerste, tweede, etc.) regularisatie is, en of daar misschien ideeën uit kunnen worden afgeleid voor alternatieve vormen die b.v. beter aansluiten op gebruikerswensen om explicieter

proces of gebiedskennis in te brengen. Gegeven een (deel)gebied ziet regularisatie er in meer algemene vorm als volgt uit.

De basis van regularisatie wordt gevormd door het gehanteerde statistische model voor de onzekerheden in de transporten. Uit dit model volgen dan de onzekerheden voor derivaten van die transporten, zoals die bijvoorbeeld voorkomen in eerste en tweede orde differenties. Hierbij moet dus worden gerealiseerd dat als men zou starten met een foutenmodel voor die differenties dat dan daar misschien geen consistent model voor de fouten in de eigenlijke transporten bij zou kunnen worden gevonden.

Voor de definitie van het foutenmodel beschouwen we voor het gemak één component T van de transporten (bijvoorbeeld T_x) en diens waarde in positie $\vec{r}_k$ wordt aangegeven met T_k . Aan deze T_k is de onzekerheid V_k toegekend. Het statistische model bestaat dan uit de aanname dat de verwachtingswaarde van V_k nul is voor elke k . De tweede orde statistiek wordt beschreven door de covariantiestructuur van de onzekerheden:

$$R_{k\ell} := E[V_k \cdot V_\ell] \quad (4.1.10)$$

Deze (symmetrische) covariantiematrix beschrijft de afhankelijkheid van de diverse onzekerheden. Alleen als alle $R_{k\ell} = 0$ ($k \neq \ell$) zijn zij onafhankelijk van elkaar en is de ruis V_k wit. In de praktijk zal de covariantiestructuur door de gebruiker moeten worden opgegeven en hier kan men, voor zover mogelijk, dan proces of gebiedskennis inbrengen. Een voorbeeld zou kunnen zijn dat men aanneemt dat de korrelatie van de onzekerheden V_k en V_ℓ afhangt van de afstand tussen de posities $\vec{r}_k$ en $\vec{r}_\ell$ waarop zij gedefinieerd zijn:

$$R_{k\ell} := R(\|\vec{r}_k - \vec{r}_\ell\|) \quad (4.1.11)$$

Overigens hoeft de te gebruiken afstandsmaat niet isotroop te zijn. In de praktijk is het zeer denkbaar dat men de wederzijdse afhankelijkheid in één bepaalde richting sterker wil kiezen dan in de richting loodrecht daarop. Bovendien hoeven deze hoofdrichtingen niet langs de x -as of de y -as te vallen. Als voorbeeld zou men die hoofdrichtingen langs en loodrecht op een geul of open modelrand kunnen nemen. Zo'n niet isotrope afstandsdefinitie ziet er als volgt uit:

$$\|\vec{r}_k - \vec{r}_\ell\|_{\Sigma^{-1}} = \sqrt{(\vec{r}_k - \vec{r}_\ell)^T \cdot \Sigma^{-1} \cdot (\vec{r}_k - \vec{r}_\ell)} \quad (4.1.12)$$

Hierin is Σ een symmetrische en positief definitieve matrix en heeft dan een zelfde vorm als een covariantiematrix. De eigenvectoren en eigenwaarden van Σ geven de hoofdrichtingen en de schaling daarlangs. Onder de afstandsmaat van vergelijking (4.1.12) hebben in 2D de isolijnen met gelijke afstand tot een punt de vorm van ellipsen die bovendien 'schuin' georiënteerd kunnen zijn.

Als een concreet voorbeeld van een spatiële korrelatiestructuur waarin de aannames van vergelijking (4.1.11) en (4.1.12) zijn gekombineerd kan dan worden genoemd:

$$R_{k\ell} := R_0 \cdot \exp(-\|\vec{r}_k - \vec{r}_\ell\|_{\Sigma^{-1}}) \quad (4.1.13)$$

Dit is een korrelatiestructuur die exponentieel met de (al dan niet isotrope) afstand afneemt. Voor punten ver van elkaar weg wordt de korrelatie zo goed als nul. Bedenk echter dat punten die onder het Σ -afstandsbegrip ver weg van elkaar liggen toch dicht bij elkaar zouden kunnen liggen bij de standaard afstandsdefinitie.

De ‘moraal’ tot nu toe is dan dat voorafgaand aan het toepassen van regularisatie er per geschikt deelgebied een model voor de onzekerheden in de te schatten transporten moet worden opgesteld. Op dat moment kunnen consistente vormen van regularisaties worden gedefinieerd. In concrete zin wordt dit nagegaan voor nulde en eerste orden van regularisatie.

Een nulde orde regularisatie bestaat er dan uit dat men binnen een $1D$ of $2D$ (deel)gebied alle of een gedeelte der T_k ($1 \leq k \leq K$) selekteert en hiervoor het volgende model voor de regularisatievergelijkingen opstelt:

$$T_k = \tilde{T}_k + V_k \quad (4.1.14)$$

De $\tilde{T}_k$ vormen een initial guess voor de te schatten T_k . Deze kunnen het resultaat zijn van metingen in strikte zin, of ‘virtuele’ metingen ontleend aan hybride of procesmodellen zoals beschreven in Hoofdstuk 2. Eventueel wordt een of ander konstante gekozen die een goede schatting voor het gemiddelde vormt. Op basis van de modelvergelijkingen en de covariantiematrix $R_{k\ell}$ van de ruizen V_k en V_ℓ moet dan de volgende kostenfunctie L_0 aan de oorspronkelijke kostenfunctie J worden toegevoegd als bijdrage van de nulde orde regularisatie:

$$L_0 = \sum_{k=1}^K \sum_{\ell=1}^K (T_k - \tilde{T}_k) \cdot R_{k\ell}^{-1} \cdot (T_\ell - \tilde{T}_\ell) \quad (4.1.15)$$

Voor de covariantiematrix zou bijvoorbeeld de vorm van vergelijking (4.1.13) kunnen worden gekozen.

Aan een *eerste* orde regularisatie ligt de aanname ten grondslag dat de transporten T_k niet of nauwelijks variëren over het (deel)gebied. Het bijbehorende model voor de transporten is dan:

$$T_k = T_0 + V_k \quad (4.1.16)$$

Hierin is T_0 het niveau van het veronderstelde konstante veld. Deze T_0 is echter niet bekend. Echter, T_0 kan worden geëlimineerd door naar verschillen $T_k - T_\ell = V_k - V_\ell$ te kijken. De covariantiestructuur van die verschillen volgt uit:

$$E[(V_k - V_\ell) \cdot (V_m - V_n)] = R_{km} - R_{kn} - R_{\ell m} + R_{\ell n} =: S_{k\ell}^{mn} \quad (4.1.17)$$

In principe kunnen dan alle mogelijke verschillen gekombineerd worden in een kostenfunctie L_1 voor de eerste orde regularisatie. Die L_1 ziet er dan als volgt uit:

$$L_1 = \sum_{k,\ell,m,n} (T_k - T_\ell) \cdot (S^{-1})_{k\ell}^{mn} \cdot (T_m - T_n) \quad (4.1.18)$$

Een speciaal geval hiervan is dat we ons net zoals in het huidige ISTM beperken tot de verschillen $T_k - T_\ell$ van naburige cellen, d.w.z. $2 \leq k \leq K$ met daarbij $\ell = k - 1$. Dan reduceert de L_1 van vergelijking (4.1.18) tot:

$$L_1 = \sum_{k,\ell=2}^K (T_k - T_{k-1}) \cdot (S^{-1})_k^\ell \cdot (T_\ell - T_{\ell-1}) \quad (4.1.19)$$

Hierin is de covariantiematrix

$$S_k^\ell := E[(V_k - V_{k-1}) \cdot (V_\ell - V_{\ell-1})] = R_{k\ell} - R_{k\ell-1} - R_{k-1\ell} + R_{k-1\ell-1} \quad (4.1.20)$$

De L_1 van vergelijking (4.1.19) is echter nog steeds algemener van vorm dan de versie zoals die in het huidige ISTM is geïmplementeerd. De laatste vormt echter voor “ $(S)_k^\ell = 0$ voor $k \neq \ell$ en $(S)_k^k$ onafhankelijk van k ” een speciaal geval van de L_1 van vergelijking (4.1.19). De in de vorige zin met “” aangegeven condities voor de $(S)_k^\ell$ kunnen worden terugvertaald naar de covariantiestructuur $R_{k\ell}$ voor de V_k . Dit leidt dan tot het model van vergelijking (4.1.9) voor de V_k en brengt ons weer naar de aldaar gegeven konklusie dat de diverse V_k maximaal gekorreleerd moeten zijn, ofwel identiek. Dat is dan alleen weer zinvol voor deelgebieden waarin een beperkte variabiliteit een redelijke aanname is.

Voor de *tweede* orde regularisatie kan een overeenkomstig betoog worden opgesteld. In dit geval is het regularisatiemodel gebaseerd op de aanname dat de transporten lineair variëren over het (deel)gebied, ofwel:

$$T_k = T_0 + \vec{\lambda}^T \cdot \vec{r}_k + V_k \quad (4.1.21)$$

Dat lineaire verloop wordt met de parameters T_0 en $\vec{\lambda}$ vastgelegd. Deze parameters zijn echter niet bekend maar kunnen met tweede orde differenties van de T_k verwijderd worden. Ook hier lijkt de algemene aanpak snel tot ingewikkelde vormen van de regularisatie componenten en (vooral) hun covariantiestructuur. Daarom lijkt een beperking tot tweede orde differenties op basis van naburige cellen zowel praktisch als conceptueel te prefereren, net zoals dat in de huidige ISTM implementatie is gedaan.

Vanuit praktisch oogpunt moet aan bovenstaande introductie van meer algemene vormen van regularisatie de volgende opmerking worden toegevoegd. Gegeven een regularisatie zoals algemeen met vergelijking (4.1.2) weergegeven, moet de oorspronkelijke kostenfunctie J worden uitgebreid tot een vorm zoals beschreven met vergelijking (4.1.3). Het blijkt dat dan de inverse R^{-1} van een covariantiematrix R moet worden berekend (zie ook de vergelijkingen 4.1.15 en 4.1.18). Dat is qua rekeninspanning een zeer intensieve en dus onaantrekkelijke taak indien de dimensie van die matrix groot is (d.w.z. veel gekoppelde regularisatievergelijkingen). Bovendien kan bij een dergelijke grote dimensie die inversie op zijn beurt een welgesteld probleem vormen. In de huidige implementatie van het ISTM hadden we daar geen last van omdat de daar toegepaste R geacht werd een diagonaalstructuur te bezitten. We hebben al gezien dat dat bepaalde consequenties had voor het feitelijke model dat dan voor de regularisatie(s) wordt gehanteerd. Uit praktijktoepassingen is echter gebleken dat dit model niet tot onzinnige resultaten leidt.

Op basis van al het voorafgaande, rekening houdende met de restrikties verwoord in de vorige alinea kunnen we dan tot de volgende uiteindelijke synthese komen:

1. De eerste en tweede orde regularisaties zoals tot nu toe toegepast kunnen worden gehandhaafd. De bevindingen van deze paragraaf als die in de volgende paragraaf suggereren echter sterk om de gekozen spreidingen in deze regularisaties niet uniform te kiezen over de het gehele rekengebied. De aanbeveling is dan om de spreidingen per deelgebied afzonderlijk in te stellen. Die deelgebieden (b.v. morfologische eenheden) moeten worden geselecteerd op basis van hun karakteristiek t.a.v. (waargenomen of verwachte) spatiële variabiliteit en/of mate van onzekerheid van de transporten aldaar, en/of geometrische eigenschappen.
2. In het vorige aandachtspunt wordt gepleit voor handhaven van de nu geïmplementeerde vormen van regularisatie. Hieraan moet wel de volgende nuancering/verfijning worden toegevoegd. De huidige 10 vormen van eerste (4) en tweede orde (6) regularisatie zijn in principe alle langs de x -as en/of de y -as gedefinieerd. Reeds op basis van fysika mag een betere aanpak worden verwacht indien meer algemeen deze regularisaties langs door de gebruiker gewenste (willekeurige) richtingen kan worden opgegeven. Dit betekent dat differenties $T_k - T_\ell$ moeten kunnen worden ingebracht waarvan de onderliggende posities $\vec{r}_k$ en $\vec{r}_\ell$ cellen vertegenwoordigen die niet strikt naast elkaar (zelfde y positie) of onder elkaar (zelfde x positie) hoeven te liggen. Bovendien moeten de componenten T_k, T_ℓ dan niet beperkt zijn tot één der T_x, T_y maar ook de component (projectie) van $\vec{T}$ in een willekeurige richting kunnen omvatten.
3. Als voorbeeld waar het zinvol is om regularisatie meer gericht dan isotroop toe te passen kunnen plaat-geulen systemen worden genoemd. Als transporten b.v. voornamelijk via een geul plaatsvinden moet de regularisatie aldaar vooral langs/longitudinaal de geul worden voorgeschreven en niet of veel minder in de transversale richting.

Bij het hier gestelde kunnen nog de volgende opmerkingen worden geplaatst:

- i. Deze funktionaliteit vereist feitelijk dat voor het regulariseren meer flexibiliteit is vereist in het voorschrijven van (lineaire) combinaties van $(T_x)_m^n$ en $(T_y)_m^n$ dan nu mogelijk is met de 10 gefixeerde vormen van differenties.
- ii. Zo'n uitbreiding is daarmee een voorbeeld van een meer generieke vorm van regularisatie.

Voor het binnen het ISTM in concrete vorm realiseren van deze flexibiliteit/genericiteit van de regularisaties, wordt voorgesteld dat de gebruiker met de invoer van een 'stempel' $\{\alpha_{\mu\nu}^x, \alpha_{\mu\nu}^y; \sigma\}$ kan specificeren dat

$$\sum_{\mu,\nu=-1}^1 \left(\alpha_{\mu\nu}^x \cdot (T_x)_{m+\mu}^{n+\nu} + \alpha_{\mu\nu}^y \cdot (T_y)_{m+\mu}^{n+\nu} \right) + V_\sigma = 0 \quad (4.1.22)$$

de gewenste regularisatievergelijking is voor cel (m,n) . Voor elke stempel moet derhalve ook aangegeven worden op welk domein die moet worden toegepast (zie de onder (1). gegeven argumenten voor spatiële variabiliteit van regularisaties). Dat domein kan een deelgebied zijn, een bepaald bepaalde traject (b.v. langs een geul), of zelfs uit één cel bestaan.

Op deze manier kunnen regularisaties gedefinieerd worden die alle nu bestaande 10 vormen omvat, maar daarnaast additionele funktionaliteit biedt. Zo kan men nu b.v. regularisatie langs willekeurige richtingen definieren, of een (diskrete) Poisson

vergelijking als regularisator opgeven, of andere vormen waarvoor al dan niet een fysische basis is aan te wijzen.

De onzekerheden in de differenties van vergelijking (4.1.22) voor verschillende cellen (m,n) worden onafhankelijk verondersteld om de inversie van grote covariantiematrices te voorkomen. Zolang in navolging van (1) de regularisaties voldoende afgestemd zijn op deelgebieden met hun specifieke ruimtelijke variabiliteit lijkt dat geen wezenlijk bezwaar of beperking.

- iii. Aan vergelijking (4.1.22) is te herkennen dat deze regularisaties *lokale* vormen zijn in de zin dat per stempel alleen de T_x en T_y componenten van naburige cellen worden gekoppeld. Deze regularisatie wordt vervolgens wel gesommeerd over alle roosterpunten (m,n) , zij het met spatieel variabele spreidingen. Conform de onder (1) gewenste niet-uniformiteit kunnen dan nog steeds deelgebieden met verschillende karakteristieken verschillend behandeld of zelfs ontkoppeld worden.

De in het voorafgaande aandachtspunt bedoelde *lokale* regularisaties worden gebiedsgewijze aan elke daarin liggende cel toegekend. Deze vormt daarmee een soort 'basis'-of 'achtergrondregularisatie' waarmee een voor dat gehele gebied gewenste (gemiddelde, of globale) sturing wordt aangegeven.

Voor bepaalde *specifieke posities en/of trajekten* binnen dergelijke gebieden (zoals geulen, of op uitgesproken overgangen tussen twee of meer gebieden) kan het wenselijk zijn om met zware regularisatie de te schatten transporten veel strikter in een bepaalde 'richting' te sturen. Dit kan weer worden gedaan met stempels zoals (4.1.22) maar met relatief kleine spreidingen voor de bijbehorende onzekerheid. Deze veel strikter gelokaliseerde en veel meer geprononceerde regularisaties worden daarmee op de 'basisregularisatie' van het vorige aandachtspunt gesuperponeerd.

- iv. Het lijkt zinvol om aan de nu geïmplementeerde *lokale* regularisaties (en/of diens generalisatie zoals hierboven voorgesteld) andere *gedistribueerde* vormen toe te voegen. Als eerste voorbeeld zou het idee gepresenteerd aan de hand van vergelijking (4.1.14) kunnen worden genoemd. Daarin zijn de $\tilde{T}_k$ een initial guess voor de te schatten T_k , zoals bijvoorbeeld 'virtuele' metingen ontleend aan hybride of procesmodellen. De posities $\tilde{r}_k$ van die $\tilde{T}_k$ hoeven niet naburig te zijn maar kunnen over een areaal verspreid liggen. Het aantal van die T_k moet weliswaar beperkt blijven opdat de benodigde inversie van een covariantiematrix niet te rekenintensief wordt.

Als tweede en nauw gerelateerd voorbeeld kan een zelfde regularisatiemodel worden gebruikt om een puntmeting over een omgeving uit te spreiden.

Op deze manieren kan wellicht meer informatie t.a.v. transportgroottes worden ingebracht om daarmee het inverse probleem beter gesteld te krijgen. Voor de covariantiematrix van de onzekerheden zal een geschikte keuze moeten worden gemaakt. Het lijkt zinvol om daarin, net zoals in het voorbeeld van vergelijking (4.1.13), afstanden en/of andere geometrische eigenschappen te betrekken.

- v. De *gedistribueerde* regularisatie van het vorige aandachtspunt (iv) kan verder worden gegeneraliseerd naar vormen waarin grootheden voorkomen die uit meerdere transportcomponenten zijn samengesteld. Als voorbeeld kan het volgende worden genoemd. Stel dat op basis van proces of gebiedskennis bekend is (of wordt vermoedt, of anderzijds binnen het ISTM wil inbrengen) hoe globaal transporten van het ene naar één of meer andere deelgebieden verloopt. Men zou dan op het niveau van die deelgebieden vergelijkingen kunnen opstellen die dergelijke kumulatieve transporten beschrijven. In deze massa-balans op gebiedsniveau zullen in de bijbehorende vergelijkingen dan sommen van transporten (al dan niet over de randen geïntegreerd) van die deelgebieden

zijn bevat. Deze vergelijkingen zijn duidelijk gekoppeld (wat in het ene gebied eruit gaat moet elders worden geborgen) en bevatten onzekerheden die eveneens afhankelijk zijn. Deze informatie kan dan als regularisatie worden toegevoegd aan een ISTM berekening om diens uitkomsten op meer globaal niveau te sturen.

4.2 Regularisatie als filter in het ISTM

Regularisatie is ingevoerd als een interpolatiemechanisme om aan de hand van gladheidseisen éénduidige transportvelden te kunnen genereren uit inkomplete meetgegevens. Hiermee werkt dit mechanisme als een soort filter op die meetdata. Tot nu toe hebben we nog niet echt vanuit een theoretisch oogpunt gekeken wat de ‘karakteristieken’ van dit filteren zijn. In deze paragraaf worden deze karakteristieken, zij het onder zekere aannames en benaderingen, in kwantitatieve zin afgeleid. Hierbij beperken we ons tot de vormen van regularisatie zoals die in het huidige ISTM zijn geïmplementeerd, en laten de meer generieke vormen van regularisatie geïntroduceerd in de vorige paragraaf buiten beschouwing. Het doel is dan om beter het effect van de tot nu toe toegepaste vormen van regularisatie in het ISTM te begrijpen zodat in praktijktoepassingen deze smoothing gericht en beter onderbouwd kan worden aangestuurd.

Om niet verzeild te raken in onnodig lange, onnodig en ondoorzichtige formulemanipulaties gaan we in deze paragraaf uit van een $1D$ omgeving met transporten gedefinieerd op een equidistant rooster. De roosterpunten geven we aan met x_k voor $1 \leq k \leq K$. Met $\Delta x := x_k - x_{k-1}$ wordt de afstand tussen opvolgende roosterpunten aangegeven, en met $T_k, \hat{T}_k$ de te berekenen en gemeten transporten te x_k . Voor het gemak laten we voor het moment gemeten richtingen en bergingen buiten beschouwing en concentreren ons op hoe met het toepassen van eerste en tweede orde regularisaties het te berekenen veld T_k volgt uit $\hat{T}_k$. Deze vraag blijkt in het spektrale domein in analytische vorm te kunnen worden opgelost en de overdrachtsfunctie voor de spektrale component met (spatiële) frequentie ν is gegeven door vergelijking (4.2.8). Deze wordt hieronder afgeleid waarna daar een aantal opmerkingen aan worden toegevoegd.

Voor het afleiden van het transportenveld T_k uit de gemeten $\hat{T}_k$ wordt het ISTM-formalisme gevolgd dat gebaseerd is op de minimalisatie van een kostenfunctie. In dit geval ziet er die kostenfunctie als volgt uit:

$$J = \sum_{k=1}^K \left(\frac{T_k - \hat{T}_k}{\sigma_T} \right)^2 + \sum_{k=2}^K \left(\frac{T_k - T_{k-1}}{\sigma_1 \cdot \Delta x} \right)^2 + \sum_{k=2}^{K-1} \left(\frac{T_{k-1} - 2T_k + T_{k+1}}{\sigma_2 \cdot \Delta x^2} \right)^2 \quad (4.2.1)$$

Hierin is σ_T de onzekerheid in de transport metingen en de σ_1, σ_2 de spreidingen voor de eerste en tweede orde regularisaties. Voor het vinden van de oplossing T_k moet deze kostenfunctie naar al die transporten geminimaliseerd worden. Hiervoor wordt voor elke T_μ $1 \leq \mu \leq K$ de afgeleide van de kostenfunctie naar deze component genomen en het resultaat op nul gesteld. Voor de transporten in het binnengebied volgt dan de volgende vergelijking:

$$\frac{T_\mu - \hat{T}_\mu}{\sigma_T^2} + \frac{-T_{\mu-1} + 2T_\mu - T_{\mu+1}}{(\sigma_1 \cdot \Delta x)^2} + \frac{T_{\mu-2} - 4T_{\mu-1} + 6T_\mu - 4T_{\mu+1} + T_{\mu+2}}{(\sigma_2 \cdot \Delta x^2)^2} = 0 \quad (4.2.2)$$

Voor de randpunten $\mu = 1, \mu = K$ gelden iets andere vergelijkingen waarvan het effect voor de oplossing in het binnengebied wordt verwaarloosd onder de aanname van een voldoende grote K . De vergelijkingen (4.2.2) laten zien dat de oplossingen T_μ lineair afhangen van de metingen $\hat{T}_\mu$ en dat de eersten dus ontstaan uit de tweeden via een lineair filter. Het effect van dat filter kan het gemakkelijkst worden afgeleid door vergelijking (4.2.2) in het spektrale domein op te lossen. Dat doen we door in vergelijking (4.2.2) voor $\hat{T}_\mu$ de volgende spektrale component in te vullen:

$$\hat{T}_\mu = \hat{T}_0 \cdot e^{i \cdot \nu \cdot \mu} \quad (4.2.3)$$

Deze vorm vertegenwoordigt een harmonische functie met periode $P = 2\pi / \nu$ gridcellen. Intuïtief is misschien duidelijker om i.p.v. ν een integer ℓ te introduceren die het aantal malen aangeeft dat de golf past op het beschouwde domein. Dan volgt de relatie tussen ℓ en ν uit $P = 2\pi / \nu = K / \ell$ zodat

$$\nu = \nu_\ell := \frac{2\pi \ell}{K} \quad (4.2.4)$$

Op deze manier geeft bij deze ν de harmonische functie van vergelijking (4.2.3) de ℓ^e boventoon aan t.o.v. de grondtoon met hoekfrequentie $\nu_1 = 2\pi / K$. Voor $\ell = 0$ is $\nu_\ell = 0$ en geeft de component van vergelijking (4.2.3) het gemiddelde van de $\hat{T}_\mu$ aan. De overige componenten met hoekfrequenties ν_ℓ beschrijven de variaties rond dat gemiddelde. Merk op dat de hoogst frequentie die zinvol op het rooster gedefinieerd is een periode omvat van twee roostercellen en voor deze $P=2$ is $\ell = \frac{1}{2} \cdot K$ is zodat de bijbehorende maximale $\nu_\ell = \pi$. Tenslotte wordt dan vermeld dat we op deze manier de metingen hebben ontbonden in een Fourier-reeks.

Om bij hoekfrequentie ν de spektrale component voor de (te identificeren) transporten te vinden wordt voor de T_μ een vorm genomen overeenkomstig met vergelijking (4.2.3):

$$T_\mu = T_0 \cdot e^{i \cdot \nu \cdot \mu} \quad (4.2.5)$$

De vormen voor $T_\mu, \hat{T}_\mu$ volgens vergelijkingen (4.2.3) en (4.2.4) worden in vergelijking (4.2.2) gesubstitueerd. Elk der drie termen wordt daarna gedeeld door de gemeenschappelijke faktor $\exp(i \cdot \nu \cdot \mu)$ hetgeen het volgende resultaat oplevert:

$$\frac{T_0 - \hat{T}_0}{\sigma_T^2} + \frac{-e^{-i\nu} + 2 - e^{i\nu}}{(\sigma_1 \cdot \Delta x)^2} T_0 + \frac{e^{-2i\nu} - 4e^{-i\nu} + 6 - 4e^{i\nu} + e^{2i\nu}}{(\sigma_2 \cdot \Delta x^2)^2} T_0 = 0 \quad (4.2.6)$$

Ofwel:

$$T_0 \left(\frac{1}{\sigma_T^2} + \frac{2 \cdot (1 - \cos(\nu))}{(\sigma_1 \cdot \Delta x)^2} + \frac{4 \cdot (1 - \cos(\nu))^2}{(\sigma_2 \cdot \Delta x^2)^2} \right) = \hat{T}_0 \left(\frac{1}{\sigma_T^2} \right) \quad (4.2.7)$$

Als resultaat wordt de volgende amplitude-overdrachtsfunctie gevonden:

$$\frac{T_0}{\hat{T}_0} = \frac{1}{1 + 2 \cdot \left(\frac{\sigma_T}{\sigma_1} \cdot \frac{1}{\Delta x} \right)^2 \cdot (1 - \cos(\nu)) + 4 \cdot \left(\frac{\sigma_T}{\sigma_2} \cdot \frac{1}{\Delta x^2} \right)^2 \cdot (1 - \cos(\nu))^2} \quad (4.2.8)$$

Bij deze afleiding en de uiteindelijke gevonden amplitude overdrachtsfunctie worden de volgende opmerkingen en kanttekeningen geplaatst.

1. In de afleiding zijn de diverse spreidingen plaats-onafhankelijk verondersteld. Indien dat niet het geval is is het niet meer mogelijk om analytisch een expliciete uitdrukking af te leiden voor de transferfunctie van amplituden. Indien men desalniettemin met plaats-afhankelijke spreidingen zou willen werken dan is het bovenstaande alleen bruikbaar voor deelgebieden waar die spreidingen inderdaad konstant zijn of weinig zullen variëren.
2. Eerder is al vermeld dat eventuele metingen van bergingen buiten beschouwing zijn gelaten. Daarnaast zijn geen richtingsvergelijkingen in de afleiding meegenomen. Voor deze omissies zijn twee redenen aan te wijzen. Op de eerste plaats is dat gedaan om via bovenstaande conceptueel helder te krijgen hoe de eerste en tweede orde regularisaties als *filter* werken op gemeten transporten. Op de tweede plaats wordt de gevonden uitdrukking van vergelijking (4.2.8) met bergingen en/of richtingen iets ingewikkelder te interpreteren t.a.v. de expliciete bijdrage van de regularisatie aan de uiteindelijke oplossing. Het weglaten van bergingen en/of richtingen betekent in feite dat aan deze grootheden een zeer grote onzekerheid is toegevoegd.
3. In bovenstaande is eveneens expliciet aangenomen dat voor elke positie x_k een gemeten transport $\hat{T}_k$ beschikbaar is. Dit volgt feitelijk al uit de eerste opmerking waarin de spreidingen σ_T konstant en eindig zijn verondersteld (terwijl aan afwezige transport-metingen formeel een onzekerheid σ_T met waarde ∞ moet worden toegekend). De situatie dat voor elk roosterpunt transportmetingen beschikbaar zijn zal in de praktijk zelden optreden. Desalniettemin kan het bovenstaande dan nog steeds goed bruikbaar zijn. Men kan bijvoorbeeld denken aan het geval dat aan de hand van een beperkt aantal metingen een first guess van het transportenveld wordt afgeleid. Dat kan op een puur handmatige manier zijn gebeurd, bijvoorbeeld aan de hand van expert-knowledge of lokale interpolaties, of met initiële en meer oriënterende ISTM berekeningen, of dat het initial guess veld afkomstig is van een hybride of procesmodel zoals in Hoofdstuk 2 beschreven. Voor al die gevallen zegt vergelijking (4.2.8) iets over het effect dat de ‘definitieve’ ISTM-run vervolgens op de (pseudo of virtuele) metingen $\hat{T}_k$ zal uitoefenen. Zie ook onderstaande aandachtspunt (iv) voor het afleiden van dergelijke pseudo of virtuele metingen $\hat{T}_k$.

Voor de interpretatie en/of praktisch gebruik van de regularisaties kan nu op basis van vergelijking (4.2.8) het volgende worden afgeleid:

- i. De regularisaties werken als een laag-doorlaat filter. Dat is te zien aan het groter worden van de noemer van vergelijking (4.2.8) voor toenemende hoekfrequentie

$\nu = \nu_\ell = 2\pi\ell / K$. De regularisatie heeft geen effect op het gemiddelde transport die in de Fourier-representatie gerepresenteerd zijn met de component $\nu = 0$ ($\ell = 0$). Het maximale effect wordt gevonden voor de Fourier component met de hoogste hoekfrequentie, namelijk $\nu = \pi$ ($\ell = \frac{1}{2}K$). De demping van deze component (en de ander niet konstante componenten) hangt af van de standaard deviaties σ_1, σ_2 toegekend aan de regularisaties.

- ii. Omdat voor de lagere frequenties $\nu \leq \frac{1}{2}\pi$ ($\ell \leq \frac{1}{4}K$) geldt dat $1 - \cos(\nu) < (1 - \cos(\nu))^2$ heeft eerste orde regularisatie daar een sterker effect dan tweede orde regularisatie. Dat is zoals verwacht omdat de eerste orde elke spatiële variatie onderdrukt, terwijl de tweede orde regularisatie ‘pas’ effect uitoefent wanneer die variaties niet-lineair zijn.
- iii. Aan vergelijking (4.2.8) is te herkennen dat het effect van de regularisaties afhangt van de verhoudingen van de standaard deviaties σ_1, σ_2 t.o.v. de spreiding σ_T van de metingen, en niet zozeer van de absolute waarden van al deze grootheden. Dit is feitelijk niets nieuws. Waar we nu wel beter inzicht in kunnen verkrijgen is hoe de σ_1, σ_2 t.o.v. elkaar gekozen moeten worden (en mede in relatie met de rooster afstand Δx) om te bewerkstelligen dat de eerste en tweede orde regularisaties in een bepaalde verhouding aanwezig zijn. Hiervoor moeten de grootte-orde van de termen in de noemer, d.w.z. 1 , $2 \cdot \left(\frac{\sigma_T}{\sigma_1} \cdot \frac{1}{\Delta x}\right)^2 \cdot (1 - \cos(\nu))$, en $4 \cdot \left(\frac{\sigma_T}{\sigma_2} \cdot \frac{1}{\Delta x^2}\right)^2 \cdot (1 - \cos(\nu))^2$ met elkaar worden vergeleken. Anderzijds, heeft men veel en betrouwbare metingen in een gebied (zodat σ_T klein is) dan mag er weinig regularisatie worden toegepast. M.b.v. bovenstaande formule kan dat ‘weinig’ gekwantificeerd worden tot de eis dat $\left(\frac{\sigma_T}{\sigma_1} \cdot \frac{1}{\Delta x}\right)^2 \ll 1$ en $\left(\frac{\sigma_T}{\sigma_2} \cdot \frac{1}{\Delta x^2}\right)^2 \ll 1$.
- iv. Als vervolg van het vorige aandachtspunt kan nu eveneens beter worden nagegaan hoe de standaard deviaties σ_1, σ_2 van de regularisaties moeten worden aangepast om bij een roosterverfijning (of vergroving) een consistente instelling van de regularisaties te verkrijgen. Dit aspect is belangrijk bij een ‘adiabatische’ schattingsprocedure. Hierbij moet men denken aan het herhaaldelijk toepassen van het ISTM waarin men in een iteratie het ISTM voedt met de uitkomsten van de vorige iteratie(s). Dat zou men kunnen doen door eerst op een grof rooster een veld af te leiden dat (al dan niet binnen de gestelde beperkingen van de ruimtelijke resolutie) aan de verwachtingen (=expert knowledge) voldoet. Naarmate het rooster grover is zal het inverse probleem beter gesteld zijn. Hierna kan op een verfijnd rooster de procedure herhaald worden. Door gebruik te maken van de resultaten van de vorige exercities (of dat gedeelte/deelgebieden waarin men vertrouwen heeft) en deze als virtuele (en eventueel geaggregeerde) metingen in te voeren kan de stabiliteit op het fijnere rooster vergroot worden. Dit kan men herhalen tot met de gewenste spatiële resolutie een bevredigende oplossing is gevonden.
- v. Aan vergelijking (4.2.8) kunnen argumenten worden ontleend om de regularisaties plaatsafhankelijk te maken. Zijn er b.v. in een deelgebied veel en nauwkeurige metingen, en/of wordt daar een grote spatiële variabiliteit verwacht dan mogen deze door de regularisaties niet te sterk onderdrukt worden. Als we ons voor het gemak even beperken tot alleen de eerste orde regularisatie en eisen dat frequenties lager dan ν_ℓ voor zekere $\ell \ll K$ met hoogstens een faktor ε , $\varepsilon \in (0,1]$ onderdrukt mogen worden dan moet gelden:

$$\frac{1}{1 + 2 \cdot \left(\frac{\sigma_T}{\sigma_1} \cdot \frac{1}{\Delta x} \right)^2 \cdot (1 - \cos(\nu_\ell))} \geq \varepsilon \rightarrow$$

$$\frac{\sigma_1}{\sigma_T} \cdot \Delta x \geq \sqrt{2 \cdot \frac{\varepsilon}{1-\varepsilon} \cdot (1 - \cos(\nu_\ell))} \approx \nu_\ell \sqrt{\frac{\varepsilon}{1-\varepsilon}} = 2\pi \cdot \frac{\ell}{K} \cdot \sqrt{\frac{\varepsilon}{1-\varepsilon}} \quad (4.2.9)$$

Met deze formule is de spreiding voor de regularisatie gekoppeld aan de de spreiding van de metingen alsmede de (verwachte) spatiële variabiliteit van het transportenveld. Daarmee volgt dan direct dat niet uniforme eigenschappen van de transportmetingen met niet uniforme regularisatie vergezeld moeten gaan. Aan de afgeleide frequentiekaracteristiek is te zien dat naarmate de regularisatie toeneemt de oplossingen gladder worden. Derhalve moet regularisatie het meest (respektievelijk minst) worden toegepast waar de laagste (respektievelijk hoogste) variabiliteit van de transportvectoren mag worden verwacht.

- vi. De vorm van laagdoorlaatfilter dat de regularisatie in het huidige ISTM vormt, en waarvan langs een 1D raai de frequentiekaracteristiek met vergelijking (4.2.8) is gegeven is het directe gevolg van de aangenomen regularisatievergelijkingen van de vorm:

$$\frac{T_k - T_{k-1}}{\Delta x} + V_1 = 0 \quad , \quad \frac{T_{k+1} - 2T_k + T_{k-1}}{\Delta x^2} + V_2 = 0$$

Op deze manier zou men voor alternatieve (lineaire) regularisatie-voorschriften het amplitude spectrum kunnen berekenen en nagaan of de dan gevonden karakteristieken beter bij de proces of gebiedskennis aansluiten. Omgekeerd, kan men echter ook uitgaan van een gewenst spectrum en dan het bijbehorende regularisatievoorschrift afleiden. Op deze manier kan men bepaalde spektrale banden explicieter/sterker dan nu onderdrukken of juist versterken. Bij het kiezen van zo'n karakteristiek zal men zich weer net als voorheen moeten laten leiden door proces of gebiedskennis.

4.3 Overige vormen voor het inbrengen van regularisatie

Additioneel aan de meer systeem-theoretische benadering van regularisatie die in de vorige paragrafen van dit hoofdstuk is gevolgd kunnen voor regularisatie/smoothing de volgende voor de praktijk relevante aspecten eveneens worden ingebracht:

1. De functie van regularisatie is het inbrengen van constraints aan de oplossing om een in eerste instantie niet goed gesteld invers probleem éénduidig te kunnen oplossen. Ons beperkend tot een lineair invers probleem betekent dit dat een stelsel $H \cdot \vec{\Theta} = \vec{Z}$ moet worden opgelost. Hierin is het aantal vergelijkingen willekeurig en kan dus net zo goed minder, gelijk of groter dan de dimensies van de gezochte oplossing $\vec{\Theta}$ zijn. Zodoende hoeft de systeemmatrix H niet noodzakelijkerwijs vierkant te zijn. In dat geval wordt zo'n stelsel in kleinste kwadraten zin opgelost, d.w.z. de $\vec{\Theta}$ moet voldoen aan $H^T H \cdot \vec{\Theta} = H^T \cdot \vec{Z}$. Als het aantal onderling onafhankelijke vergelijkingen kleiner is dan de dimensie van $\vec{\Theta}$ dan is dit stelsel echter singulier en niet éénduidig oplosbaar. Om toch een eenduidige oplossing te genereren wordt dan vaak de $\vec{\Theta}$ gekozen met de kleinste norm (ofwel lengte). Het blijkt dat deze heel elegant kan worden gevonden aan de hand van een zogenaamde singuliere waarden decompositie (SVD) van H (zie b.v. Press et al., 1994). Deze weg zou ook voor het vinden van éénduidige ISTM-oplossingen

kunnen worden gevolgd ware het niet dat in de praktijk de dimensies van de matrix H en vektor $\bar{\Theta}$ meestal te groot is om qua rekeninspanning en geheugengebruik zo'n SVD echt te kunnen uitvoeren. Het blijkt echter dat de oplossing $\bar{\Theta}_0$ met een minimale norm ook als limiet van een reeks $\bar{\Theta}_\varepsilon$, $\varepsilon \downarrow 0$, kan worden gevonden waarbij $\bar{\Theta}_\varepsilon$ de oplossing is van het minimalisatieprobleem $\text{Min}_{\bar{\Theta}} \left(\|H \cdot \bar{\Theta} - \bar{Z}\|^2 + \varepsilon \cdot \|\bar{\Theta}\|^2 \right)$. Merk op dat in deze zin de term $\varepsilon \cdot \|\bar{\Theta}\|^2$ eveneens als een regularisatie moet worden geïnterpreteerd.

De procedure om via een limiet $\bar{\Theta}_\varepsilon$, $\varepsilon \downarrow 0$, de oplossing (met minimale norm) van een slecht gekonditioneerd invers probleem te vinden is dan als volgt. Voor een geschikte $\varepsilon_1 > 0$ wordt (met een iteratieve methode, en met initial guess $\bar{\Theta}_{ini} = \bar{0}$; net zoals dat nu binnen het ISTM wordt gedaan) de oplossing $\bar{\Theta}_1$ van het minimalisatieprobleem $\text{Min}_{\bar{\Theta}} \left(\|H \cdot \bar{\Theta} - \bar{Z}\|^2 + \varepsilon_1 \cdot \|\bar{\Theta}\|^2 \right)$ gezocht. Een "geschikte $\varepsilon_1 > 0$ " betekent dat ε_1 groot genoeg moet zijn om een éénduidige $\bar{\Theta}_1$ te vinden. Dit is altijd mogelijk omdat voor grote ε_1 de $\|\bar{\Theta}\|^2$ -term in de kostenfunctie dominant is. Vervolgens wordt voor een $\varepsilon_2 < \varepsilon_1$ dit recept herhaald maar nu met $\bar{\Theta}_1$ als initial guess om een $\bar{\Theta}_2$ als update voor de oplossing te vinden. Voor een $\varepsilon_3 < \varepsilon_2$ wordt deze $\bar{\Theta}_2$ als initial guess gebruikt om een $\bar{\Theta}_3$ te vinden. Door verdere herhalingen wordt voor $\varepsilon_n \downarrow 0$ de gewenste oplossing gevonden.

De essentie van de aanpak is dat voor het oorspronkelijke niet goed gestelde probleem toch een éénduidige oplossing wordt gevonden door de oplossingsprocedure met een bepaalde conditie te sturen. Hier bestond die conditie uit het minimaal zijn van de norm, ofwel de som van de kwadraten van de componenten van $\bar{\Theta}$.

Zo'n 'adiabatische' aanpak zou ook voor het ISTM bruikbaar kunnen zijn om bij een slecht gesteld probleem (hetgeen in de praktijk bij weinig tot geen informatie over transportgrootten vrijwel altijd het geval is) toch éénduidige, zinvolle, en vooral stuurbare oplossingen te kunnen genereren. Het bovenstaande om een minimum-norm-oplossing te vinden zou een redelijk alternatief kunnen zijn. Anderzijds zouden voor die sturing naar éénduidige oplossingen ook andere regularisatievormen (of combinaties daarvan) zoals hierboven beschouwd zinvol kunnen zijn.

In zo'n aanpak wordt dan een serie van oplossingen $\bar{\Theta}_n$ gevonden. Het ligt dan voor de hand dat de gebruiker het verloop van die oplossingen in nader detail evalueert op hun realisme. Dit zou tot ingrepen in het iteratie proces kunnen leiden. Zo kunnen bijvoorbeeld voor gebieden waar na enkele iteraties een goede performance wordt gevonden de daar gevonden transporten als virtuele metingen in de opvolgende iteraties kunnen worden opgenomen. Anderzijds kunnen tussentijds de regularisaties die niet in het 'relaxatie'-mechanisme zijn opgenomen zo nodig qua grootte, variabiliteit of zelfs qua vorm nader aangepast worden. Op deze en andere manier kan de gebruiker dan op basis van zijn expertise het inverse proces nog explicieter sturen.

2. Een variant van het voorafgaande aandachtspunt zou kunnen bestaan uit het in eerste instantie toepassen van het ISTM op een relatief grof rooster, of zelfs op basis van een partitie van het rekendomein in deelgebieden. Op zo'n grotere schaal is het inverse probleem beter gekonditioneerd. Bovendien kunnen dan relatief snel diverse oriënterende berekeningen worden uitgevoerd die al een goed idee geven voor trends in de oplossingen en hun gevoeligheid voor variaties in de diverse vormen van

regularisatie. Realistisch ogende resultaten kunnen vervolgens als virtuele metingen (mogelijk in de vorm van geaggregeerde grootheden) worden toegevoegd bij latere berekeningen op fijnere roosters.

3. Alle vormen van regularisatie zoals die tot nu toe zijn beschouwd hebben betrekking op het smoothen/filteren/interpoleren van transporten. In ISTM toepassingen in de praktijk vormen echter de gemeten bergingen de voornaamste bron van meetinformatie, zij het in indirecte vorm. Onzekerheden in die bergingen worden weliswaar meegenomen in de door het ISTM uitgevoerde integratie maar het is niet geheel duidelijk in welke mate deze het uiteindelijke resultaat beïnvloeden. Dit dient in nader detail onderzocht te worden. Een eerste inventarisatie van ideeën levert het volgende op. Op de eerste plaats zouden aan de hand van een praktisch case gevoeligheidsanalyses kunnen worden uitgevoerd voor parameter variaties in het model voor de onzekerheden in de bergingen. Merk hierbij op dat vanuit theoretisch oogpunt (continuïteitsvergelijking) de onzekerheden in transporten en bergingen niet geheel los van elkaar kunnen worden gezien. Uit onnauwkeurige bergingen mogen dan bijvoorbeeld geen nauwkeurige schattingen voor transporten worden verwacht.

Op dit moment zijn in het ISTM de onzekerheden in de bergingen gemodelleerd met een spatueel witte ruis. De enige varieerbare parameter is de spreiding van die ruis. Een volgende stap zou kunnen zijn dat rekening wordt gehouden met een spatiële korrelatie van de onzekerheden in de bergingen.

Indien een aanzienlijke onzekerheid in de bergingen aanwezig is kan het praktisch gesproken effectiever zijn om deze voorafgaand aan een ISTM-run via pre-filtering te reduceren. Dit filter moet dan gerelateerd worden met het gehanteerde model voor de onzekerheden, en hoeft daarmee niet spatueel uniform te zijn. De karakteristieken van dit filter kunnen worden afgeleid op de wijze waarop dat in Paragraaf 4.2 is gedaan voor het vinden van de filterkarakteristiek van de op dit moment in het ISTM gebruikte regularisatie-mechanismen.

4.4 Samenvatting aanbevelingen regularisatie

Op basis van de in de vorige paragrafen van dit hoofdstuk gepresenteerde achtergrond en ideeën t.a.v. regularisatie kunnen als samenvatting de volgende voor de praktijk relevante aanbevelingen worden opgesteld. Deze aanbevelingen zijn alle van vergelijkbare prioriteit zodat de gegeven volgorde als min of meer willekeurig kan worden beschouwd.

1. Niet uniforme regularisatie biedt voordelen t.o.v. uniforme regularisatie. Door de regularisaties spatueel afhankelijk te maken (per deelgebieden, zeg morfologische eenheden) kan deze beter kan worden afgestemd op de lokale karakteristieken van de transporten zoals ruimtelijke variabiliteit en onzekerheid.
2. De nu in het ISTM gehanteerde 10 vormen van *lokale* regularisatie omvatten eerste en tweede orde differenties van transporten van naburige cellen. Deze lokale differenties worden langs de x of y richtingen genomen en op alle (m,n) -cellen toegepast. Het is zinvol om deze functionaliteit uit te breiden naar een veel meer generieke vorm van lokale differenties zoals concreet met de 'stempels' van vergelijking (4.1.22) is gespecificeerd. Conform (1) moeten de spreidingen in deze meer generieke vormen plaatsafhankelijk en/of per deelgebied kunnen worden opgegeven.
3. Additioneel aan de 'basis'-of 'achtergrondregularisatie' van punt (2) die gebiedsgewijze aan elke daarin liggende cel wordt toegekend kunnen voor bepaalde specifieke posities (en/of trajekten) binnen dergelijke gebieden de transporten (nog) meer expliciet in een bepaalde 'richting' worden gestuurd door op die lokaties extra zware regularisatievormen toe te voegen.

4. Naast de *lokale* regularisaties van de vorige twee aandachtspunten wordt aanbevolen om eveneens meer *gedistribueerde* vormen van regularisatie toe te passen. Op deze wijze kan informatie van transportgroottes over een deelgebied gedistribueerd worden en daarmee het inverse probleem mogelijk beter gesteld. In dit feitelijke interpoleren kunnen afstanden en/of andere geometrische eigenschappen worden betrokken.
5. De *gedistribueerde* regularisatie van aandachtspunt 4 kan verder worden gegeneraliseerd naar vormen waarin grootheden voortkomen die uit meerdere transportcomponenten zijn samengesteld. Als voorbeeld kan worden gedacht aan het opstellen van (massa)balansen op gebiedsniveau. De vergelijkingen op dit niveau, en/of smoothings-termen uitgedrukt in de daarin voorkomende makro-transporten kunnen dan als regularisatie worden toegevoegd om een ISTM berekening ook voor grotere spatiële schalen expliciet te sturen.
6. Singuliere lineaire stelsels kunnen in een gegeneraliseerde vorm éénduidig worden opgelost met behulp van een singuliere waarde dekompositie (SVD). In de praktijk is het stelsel vergelijkingen in een ISTM run meestal veel te groot om zo'n SVD-procedure uit te kunnen voeren. In dat geval kan het zelfde resultaat via een iteratieve procedure worden benaderd. Voor het ISTM wordt aanbevolen om dit aan de hand van een geschikte praktijk-case te testen en de resultaten te vergelijken met de oplossingen die met de nu geïmplementeerde solvers wordt gevonden.
7. Aanbevolen wordt om te onderzoeken of het eerst oplossen van het ISTM op een grof rooster, en de gevonden resultaten dan (geheel of gedeeltelijk) als invoer te gebruiken in opvolgende berekeningen op verfijnde roosters, perspectieven kan bieden voor het verbeteren van de robuustheid van het te vinden transporten- en bergingenveld.
8. Er wordt aanbevolen om in nader detail na te gaan in hoeverre onzekerheden in de gemeten bergingen de resultaten van een ISTM-run beïnvloeden. In het bijzonder moeten de perspectieven worden onderzocht van het a priori door filteren/smoothen reduceren van de onzekerheden in die bergingen

5 Opzet van het niet-lineaire ISTM

Om duidelijk te maken hoe het lineaire ISTM (Koster en Van den Boogaard, 1993; Van den Boogaard en Mulder, 1997) een basis vormt voor de uitbreiding naar een niet-lineair ISTM wordt eerst in Paragraaf 5.1 het lineaire ISTM samengevat, en getoond hoe het oplossen van dit lineaire stochastische systeem kan worden vertaald naar een minimalisatieprobleem. Aan de hand van deze vertaling wordt in Paragraaf 5.2 dan het concept van een niet-lineair ISTM gedefinieerd. Dit concept wordt vervolgens in Hoofdstuk 6 uitgewerkt voor het verwerken van niet-lineaire richtingsvergelijkingen, welke daadwerkelijk zijn geïmplementeerd in versie 2.00 van het ISTM.

5.1 Review van het lineaire ISTM model

Zonder de vergelijkingen in het lineaire ISTM (versie 1.03) expliciet te vernoemen kan het mathematisch/numerieke model en de oplosprocedure als volgt worden samengevat. Gegeven de meetinformatie in de vorm van bergingen, transporten, en transportrichtingen, alsmede de door de gebruiker geactiveerde regularisaties, bestaat het conceptuele model uit een stelsel stochastische lineaire vergelijkingen van de vorm:

$$H \cdot \vec{\Theta} + \vec{V} = \vec{Z} \quad (5.1.1)$$

Hierin is $\vec{\Theta}$ een vektor die wordt verkregen door alle transportcomponenten $(T_x)_m^n, (T_y)_m^n$ in een of andere volgorde onder elkaar te zetten. De dimensie van die vektor wordt aangegeven met K , zodat $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3, \dots, \Theta_K)^T$. De gekozen administratie bepaalt met welke $(T_x)_m^n$ of $(T_y)_m^n$ de component Θ_k overeenkomt en vice versa.

De matrix H bevat de coëfficiënten van alle lineaire vergelijkingen die in een ISTM toepassing gedefinieerd zijn. Het aantal van die vergelijkingen wordt aangegeven met L en dit aantal hangt af van de hoeveelheid puntmetingen van bergingen, transporten, richtingen, en de geactiveerde regularisaties. Matrix H is dan van dimensie $L \times K$. Vergelijking (5.1.1) kan alleen zinvol worden opgelost als het aantal vergelijkingen minstens zo groot, maar liever veel groter, is dan het aantal componenten van de transport vektor. Derhalve $L \gg K$.

De rechterlid vektor $\vec{Z}$ bevat de metingen in strikte zin, of wanneer een modelvergelijking van een regularisatie afkomstig is de 'metingen' in virtuele zin. In dat geval is de bijbehorende component van het rechterlid $\vec{Z}$ gelijk aan nul.

De vektor $\vec{V}$ van dimensie L representeert de modelonzekerheden. Deze onzekerheden zijn in statistische zin gedefinieerd in de vorm van hun gemiddelde en covariantiematrix. Er wordt aangenomen dat er geen systematische fouten zijn zodat het gemiddelde $E[\vec{V}] = \vec{0}$.

Hier is $E[\cdot]$ de verwachtingswaarde operator. De $L \times L$ covariantiematrix is dan gedefinieerd door $R_v := E[\vec{V} \cdot \vec{V}^T]$.

Het lineaire stelsel van vergelijking (5.1.1) kan in Maximum Likelihood worden opgelost met als resultaat de volgende optimale schatter voor de transportvektor:

$$\hat{\Theta} = (H^T \cdot R_V^{-1} \cdot H)^{-1} \cdot H^T \cdot R_V^{-1} \cdot \bar{Z} \quad (5.1.2)$$

Numeriek wordt daarbij niet de vorm van vergelijking (5.1.2) aangehouden aangezien de berekening van de inverse van de (in praktijktoepassingen vaak zeer grote) $K \times K$ matrix $H^T \cdot R_V^{-1} \cdot H$ praktisch niet haalbaar is. Deze inversie hoeft feitelijk ook niet te worden uitgevoerd aangezien $\hat{\Theta}$ gemakkelijker kan worden gevonden door het oplossen van het lineaire stelsel:

$$(H^T \cdot R_V^{-1} \cdot H) \cdot \hat{\Theta} = H^T \cdot R_V^{-1} \cdot \bar{Z} \quad (5.1.3)$$

De basis voor een uitbreiding van het ISTM naar een niet-lineaire versie wordt gevormd door het feit dat de oplossing van vergelijking (5.1.2) ook kan worden gevonden door uit het conceptuele model van vergelijking (5.1.1) een optimalisatieprobleem af te leiden. Het blijkt namelijk dat de $\hat{\Theta}$ van vergelijking (5.1.2) de oplossing is van het volgende minimalisatie probleem:

$$\underset{\bar{\Theta}}{\text{Min}} (H \cdot \bar{\Theta} - \bar{Z})^T \cdot R_V^{-1} \cdot (H \cdot \bar{\Theta} - \bar{Z}) \quad (5.1.4)$$

Dit kan worden ingezien door in het Maximum Likelihood formalisme, en onder aanname van Gaussische verdelingen van de ruizen, de (minus) LogLikelihood functie te bepalen. Dan wordt afgezien van enkele termen die onafhankelijk van $\bar{\Theta}$ zijn de kwadratische functie in het rechterlid van (5.1.4) gevonden. De Maximum Likelihood schatter van $\bar{\Theta}$ is dan precies de $\hat{\Theta}$ waarvoor deze minimaal is. Anderzijds kan van de kwadratische functie in vergelijking (5.1.4) relatief gemakkelijk in analytische zin het minimum worden bepaald (namelijk via de gradiënt naar $\bar{\Theta}$) om dan de $\hat{\Theta}$ van vergelijking (5.1.2) te vinden.

De moraal is dus dat de Maximum Likelihood oplossing van vergelijking (5.1.1) equivalent is met het kwadratische minimalisatie probleem van vergelijking (5.1.4). Om dit beter inzichtelijk te maken herhalen we het bovenstaande voor het speciale geval dat de onzekerheden van verschillende modelvergelijkingen ongecorreleerd (of zelfs onafhankelijk) zijn. Dan geldt dat de $L \times L$ covariantiematrix $R_V := E[\bar{V} \cdot \bar{V}^T]$ een diagonaal structuur heeft. In entry $(R_V)_{\ell, \ell}$ is dan de variantie σ_ℓ^2 van de ruis V_ℓ te vinden die werd toegevoegd aan de ℓ^e modelvergelijking. Deze variantie is het kwadraat van de spreiding σ_ℓ . Die ℓ^e vergelijking in de vektorvorm van vergelijking (5.1.1) ziet er in uitgeschreven vorm dan als volgt uit:

$$\sum_{k=1}^K H_{\ell, k} \cdot \Theta_k + V_\ell = Z_\ell \quad , \quad 1 \leq \ell \leq L \quad (5.1.5)$$

Alle termen in deze vergelijking worden nu geschaald met de spreiding σ_ℓ van de ruisterm V_ℓ zodat:

$$\sum_{k=1}^K \frac{H_{\ell, k}}{\sigma_\ell} \cdot \Theta_k + \frac{V_\ell}{\sigma_\ell} = \frac{Z_\ell}{\sigma_\ell} \quad , \quad 1 \leq \ell \leq L \quad (5.1.6)$$

ofwel:

$$\sum_{k=1}^K \tilde{H}_{\ell,k} \cdot \Theta_k + \tilde{V}_\ell = \tilde{Z}_\ell \quad , \quad 1 \leq \ell \leq L \quad (5.1.7)$$

waarbij

$$\tilde{H}_{\ell,k} := \frac{H_{\ell,k}}{\sigma_\ell} \quad , \quad \tilde{V}_\ell := \frac{V_\ell}{\sigma_\ell} \quad , \quad \tilde{Z}_\ell := \frac{Z_\ell}{\sigma_\ell} \quad (5.1.8)$$

Hierin zijn al de ruizen $\tilde{V}_\ell$ onafhankelijk van elkaar en van gelijke variantie ter grootte 1. De covariantiematrix $R_{\tilde{V}}$ van $\tilde{V}_\ell$ is dan de eenheidsmatrix. Met $R_{\tilde{V}} = I$ is de oplossing $\hat{\Theta}$ van (5.1.7) in Maximum Likelihood zin dan analoog aan vergelijking (5.1.2) gegeven door:

$$\hat{\Theta} = (\tilde{H}^T \cdot \tilde{H})^{-1} \cdot \tilde{H}^T \cdot \tilde{Z} \quad (5.1.9)$$

Deze vergelijking representeert dus het speciale geval van vergelijking (5.1.2) bij de aanname van de onafhankelijkheid van de ruizen V_ℓ . Deze konversie kan eveneens worden doorgevoerd in de vertaling van de Maximum Likelihood schatting naar een minimalisatie-probleem. De met vergelijking (5.1.4) overeenkomstige uitdrukking wordt dan:

$$\underset{\hat{\Theta}}{\text{Min}} \left(\tilde{H} \cdot \tilde{\Theta} - \tilde{Z} \right)^T \cdot \left(\tilde{H} \cdot \tilde{\Theta} - \tilde{Z} \right) \quad (5.1.10)$$

Ook deze vorm kan gemakkelijk met analytische middelen geminimaliseerd worden om voor $\hat{\Theta}$ het resultaat van vergelijking (5.1.9) te vinden. Alvorens aan deze ‘nieuwe’ vormen een interpretatie toe te voegen wordt de vektor vorm in vergelijking (5.1.10) herschreven in een vorm die de individuele vergelijkingen bevat. Met de definitie van de volgende kostenfunctie $J(\cdot)$,

$$J(\Theta) := \left(\tilde{H} \cdot \tilde{\Theta} - \tilde{Z} \right)^T \cdot \left(\tilde{H} \cdot \tilde{\Theta} - \tilde{Z} \right) \quad (5.1.11)$$

is

$$\begin{aligned} J(\Theta) &= \sum_{\ell=1}^L \left(\sum_{k=1}^K \tilde{H}_{\ell,k} \cdot \Theta_k - \tilde{Z}_\ell \right)^2 \\ &= \sum_{\ell=1}^L \left(\frac{\sum_{k=1}^K H_{\ell,k} \cdot \Theta_k - Z_\ell}{\sigma_\ell} \right)^2 \end{aligned} \quad (5.1.12)$$

Deze transitie van een Maximum Likelihood formalisme naar een minimalisatie van een kwadratische kostenfunctie is binnen het ISTM kader niet iets nieuws. Deze vormde namelijk de basis van de definitie van de diagnostische tool, zie Appendix A en B in Van den Boogaard (1994). Daarin werd door in de sommatie over ℓ (d.w.z. de modelvergelijkingen) onderscheid te maken of de betreffende vergelijking betrekking heeft op metingen van bergingen (kv : kontinuïteitsvergelijkingen), of betrekking heeft op metingen van transporten (mv : meetvergelijkingen), of betrekking heeft op eerste (reg_1) of

tweede orde (reg_2) regularisaties, een partitie van de kostenfunctie aangebracht over deeltkostenfuncties t.a.v. de hierboven genoemde herkomst. Ofwel:

$$J_{ISTM} = J_{kv} + J_{mv} + J_{reg_1} + J_{reg_2} \quad (5.1.13)$$

Bij de vergelijkingen (5.1.12) en (5.1.13) moet wel worden opgemerkt dat zij essentieel voortkomen uit de aanname van onafhankelijkheid van de diverse ruizen. In dat geval is de kostenfunctie (of diens componenten) dan als volgt opgebouwd. Elk der ISTM vergelijkingen wordt eerst herleid in een vorm met 0 in het rechterlid. Van het resulterende linkerlid wordt het deterministische deel genomen (ofwel de modelvergelijkingen wanneer deze nog niet in een stochastische omgeving ingebed zijn) en deze wordt vervolgens gedeeld door de spreiding van de ruisterm die de onzekerheid van dat deterministische deel aangeeft. Van al deze zo gevonden termen wordt het kwadraat genomen waarna zij worden opgeteld.

5.2 Definitie van het niet-lineaire ISTM

Het recept aan het einde van de vorige paragraaf t.a.v. het definiëren van een kostenfunctie vormt de basis voor een niet-lineair ISTM. Immers wordt het lineaire ISTM uitgebreid met vergelijkingen die niet-lineair zijn in de transporten dan worden deze zoals voorheen is beschreven eerst tot nul herleid, daarna worden de linkerleden gedeeld door de spreiding van de onzekerheid die aan het linkerlid wordt toegekend, waarna het kwadraat van de zo gevonden uitdrukking aan de kostenfunctie wordt toegevoegd. Om dit met de bijbehorende formules expliciet te maken wordt de ℓ^e (en mogelijk niet-lineaire) ISTM-vergelijking weergegeven met

$$H_\ell(\vec{\Theta}) = 0 \quad (5.2.1)$$

Dit is dan de deterministische vorm van de ℓ^e vergelijking. De bijbehorende stochastische vergelijking is

$$H_\ell(\vec{\Theta}) = V_\ell \quad (5.2.2)$$

waarbij σ_ℓ de spreiding voor de onzekerheid V_ℓ van de uitdrukking $H_\ell(\vec{\Theta})$ is. De kostenfunctie van een niet-lineair ISTM is dan als volgt gedefinieerd:

$$J(\vec{\Theta}) = \sum_{\ell=1}^L \left(\frac{H_\ell(\vec{\Theta})}{\sigma_\ell} \right)^2 = \sum_{\ell=1}^L \left(\tilde{H}_\ell(\vec{\Theta}) \right)^2 \quad (5.2.3)$$

Als voorbeeld van één zo'n $H_\ell(\cdot)$ zou de continuïteitsvergelijking voor een cel (m, n) kunnen worden genoemd. Met een administratie van de vorm:

$$\Theta_{k_1} = (T_x)_m^n, \quad \Theta_{k_2} = (T_x)_{m+1}^n, \quad \Theta_{k_3} = (T_y)_m^n, \quad \Theta_{k_4} = (T_y)_m^{n+1} \quad (5.2.4)$$

is dan

$$H_\ell(\vec{\Theta}) = \frac{\Theta_{k_2} - \Theta_{k_1}}{\Delta x} + \frac{\Theta_{k_4} - \Theta_{k_3}}{\Delta y} + B_m^n \quad (5.2.5)$$

Deze $H_\ell(\cdot)$ is ‘toevallig’ nog lineair in het argument $\vec{\Theta}$ maar het idee blijft hetzelfde. Voor een echte niet-lineaire vorm wordt verwezen naar de richtingsvergelijkingen afgeleid in Hoofdstuk 6.

De kostenfunctie van vergelijking (5.2.3) representeert dan nog steeds een least squares probleem maar nu is dat mogelijk een *non-linear least squares criterium*, en niet meer een *linear least squares criterium* zoals voorheen. Voor het vinden van het optimale transportvektorveld $\vec{\Theta}$ moet de kostenfunctie van vergelijking (5.2.3) naar die $\vec{\Theta}$ geminimaliseerd worden. Vanwege de niet-lineaire modelvergelijkingen kan voor de gezochte $\hat{\Theta}$ niet meer direct een analytische uitdrukking worden gegeven en moeten numerieke minimalisatieprocedures worden aangewend om (iteratief) de $\hat{\Theta}$ te vinden. Het meest efficiënt hiervoor zijn methoden gebaseerd op de gradiënt van de kostenfunctie. Deze gradiënt is gedefinieerd door:

$$\vec{\nabla}_{\vec{\Theta}} J := \left(\frac{\partial J}{\partial \Theta_1}, \frac{\partial J}{\partial \Theta_2}, \frac{\partial J}{\partial \Theta_3}, \dots, \frac{\partial J}{\partial \Theta_K} \right)^T \quad (5.2.6)$$

Meer concreet zijn voor de kostenfunctie van vergelijking (5.2.1) de componenten van die gradiënt dan gegeven door:

$$\frac{\partial J}{\partial \Theta_k} = 2 \sum_{\ell=1}^L \frac{H_\ell(\vec{\Theta})}{\sigma_\ell^2} \cdot \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_k} = 2 \sum_{\ell=1}^L \tilde{H}_\ell(\vec{\Theta}) \cdot \frac{\partial \tilde{H}_\ell(\vec{\Theta})}{\partial \Theta_k} \quad (5.2.7)$$

Wanneer we als illustratie het $H_\ell(\cdot)$ -voorbeeld van vergelijking (5.2.4) en (5.2.5) beschouwen dan zijn de in vergelijking (5.2.7) voorkomende afgeleiden $\frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_k}$ van de volgende vorm:

$$\begin{cases} \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_{k_1}} = -\frac{1}{\Delta x}, \quad \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_{k_2}} = \frac{1}{\Delta x}, \quad \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_{k_3}} = -\frac{1}{\Delta y}, \quad \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_{k_4}} = \frac{1}{\Delta y}, \\ \frac{\partial H_\ell(\vec{\Theta})}{\partial \Theta_k} = 0 \quad \text{voor } k \neq k_1, k_2, k_3, k_4 \end{cases} \quad (5.2.8)$$

Terugkerende naar de feitelijke minimalisatie van de kostenfunctie van vergelijking (5.2.3) en het daarbij gebruiken van gradiëntmethoden wordt bij een eerste schatting $\vec{\Theta}_0$ de gradiënt berekend voor die positie. Voor die eerste schatting van $\vec{\Theta}$ zou bijvoorbeeld de $\hat{\Theta}$ van het huidige lineaire ISTM kunnen worden genomen. Vanuit deze of een andere eerste schatting en de daar berekende gradiënt wordt een verbeterde (d.w.z. met een lagere kostenfunctie) schatting $\vec{\Theta}_1$ afgeleid. Dit spel wordt herhaald leidende tot opvolgende en steeds betere schattingen $\vec{\Theta}_2, \vec{\Theta}_3, \vec{\Theta}_4, \dots$, tot dat geen (‘echte’) verbeteringen meer worden gevonden. In deze iteratieve procedure kan telkens informatie van een of meerdere eerdere stappen worden gebruikt om die convergentie zo efficiënt mogelijk te maken. Hoe dat dan wordt gedaan hangt van de gekozen gradiënt-minimalisatie-methode af. Geschikte

kandidaten daarvoor zijn Geconjugeerde Gradienten (CG) of (limited memory) Quasi Newton (QN) methoden. Zie bijvoorbeeld Press et al. (1994).

Afgezien van de feitelijk schatting van de optimale $\hat{\Theta}$ moet nu nog worden nagegaan welke onzekerheid hieraan is toegevoegd. Die onzekerheid wordt weergegeven met de covariantiematrix P van de oplossing $\hat{\Theta}$. Voor het lineaire model samengevat met de vergelijkingen (5.1.7) tot en met (5.1.12) werd gevonden dat de *inverse* van die covariantiematrix voldoet aan:

$$P^{-1} = \tilde{H}^T \cdot \tilde{H} \quad (5.2.9)$$

Zie Koster en Van den Boogaard (1993ab). Deze vorm volgde uit het Maximum Likelihood formalisme maar kan ook binnen de hier geherdefinieerde optimalisatieomgeving worden afgeleid. De matrix P^{-1} is dan gerelateerd met de Hessiaan van de kostenfunctie J van vergelijking (5.1.11). De Hessiaan van een functie is hierbij de matrix van tweede orde afgeleides van die functie naar diens argumenten en in het ISTM kader is dan:

$$(P^{-1})_{k,\ell} = \frac{1}{2} \frac{\partial^2 J}{\partial \Theta_k \partial \Theta_\ell} \bigg|_{\Theta=\hat{\Theta}} \quad (5.2.10)$$

Door de vorm van de J van vergelijking (5.1.11) in het rechterlid te substitueren kan de vergelijking (5.2.9) gemakkelijk worden afgeleid. De definitie van de covariantiematrix volgens vergelijking (5.2.10) wordt nu ook gehanteerd voor het niet-lineaire ISTM.

Met de niet-lineaire $\tilde{H}_\ell(\cdot)$ componenten in de kostenfunctie van vergelijking (5.2.3) wordt dan de volgende uitdrukking voor de (μ, ν) -entry van matrix P^{-1} gevonden:

$$(P^{-1})_{\mu,\nu} = \sum_{\ell=1}^L \left(\frac{\partial^2 \tilde{H}_\ell}{\partial \Theta_\mu \partial \Theta_\nu} \cdot \tilde{H}_\ell + \frac{\partial \tilde{H}_\ell}{\partial \Theta_\mu} \cdot \frac{\partial \tilde{H}_\ell}{\partial \Theta_\nu} \right) \bigg|_{\Theta=\hat{\Theta}} \quad (5.2.11)$$

Merk op dat als de $\tilde{H}_\ell(\cdot)$ lineair in $\tilde{\Theta}$ zijn dat dan de term in (5.2.11) met de tweede afgeleiden verdwijnt en met $\frac{\partial \tilde{H}_\ell}{\partial \Theta_\mu} = \tilde{H}_{\ell,\mu}$ reduceert bovenstaande vergelijking voor P^{-1} weer naar de vorm (5.2.9) voor het lineaire ISTM. Dit feit zou kunnen worden gebruikt voor het geval dat binnen de oplossingsprocedure onderscheid wordt gemaakt naar enerzijds de lineaire vergelijkingen (zeg: de vergelijkingen met label $1 \leq \ell \leq l$) en anderzijds de niet-lineaire vergelijkingen (zeg: de vergelijkingen met label $l < \ell \leq L$). In dat geval wordt vergelijking (5.2.11) van de vorm:

$$(P^{-1})_{\mu,\nu} = (\tilde{H}_*^T \cdot \tilde{H}_*)_{\mu,\nu} + \sum_{\ell=l+1}^L \left(\frac{\partial^2 \tilde{H}_\ell}{\partial \Theta_\mu \partial \Theta_\nu} \cdot \tilde{H}_\ell + \frac{\partial \tilde{H}_\ell}{\partial \Theta_\mu} \cdot \frac{\partial \tilde{H}_\ell}{\partial \Theta_\nu} \right) \bigg|_{\Theta=\hat{\Theta}} \quad (5.2.12)$$

waarbij de matrix $\tilde{H}_*$ betrekking heeft op het lineaire sub stelsel.

Het van het lineaire ISTM bekende probleem dat we de inverse covariantiematrix in handen krijgen en niet zozeer de gewenste covariantiematrix zelf blijft met het voorafgaande

eveneens aktueel binnen het niet-lineaire ISTM. Gegeven de omvang van de dimensie van $\bar{\Theta}$ in praktijktoepassingen is het (nog steeds) niet mogelijk om de noodzakelijke matrix inversie uit te voeren. Vooralsnog houden we de door Koster en Van den Boogaard (1993b) voorgestelde approximatie aan dat van P^{-1} de diagonaal elementen $A_{\mu,\mu}$ worden bepaald en dat dan voor de variantie van de geschatte transport component $\hat{\Theta}_{\mu}$ de benadering $1/A_{\mu,\mu}$ wordt aangehouden.

Hiermee is het concept van een niet-lineair ISTM en de eventueel te hanteren oplosprocedure vastgelegd. Samenvattend kan worden gezegd dat de equivalentie van het in het oorspronkelijke ISTM gehanteerde Maximum Likelihood principe met een minimalisatie-probleem de basis van deze niet-lineaire uitbreiding vormt.

6 Transportrichtingen in het niet-lineaire ISTM

In dit hoofdstuk wordt beschreven hoe in een niet-lineair ISTM richtingsinformatie van transporten kan worden ingebracht die beter voldoet dan de wijze waarop dat tot nu toe in het lineaire model is gerealiseerd. Op basis van het concept van een niet-lineair ISTM zoals beschreven in Hoofdstuk 5 moet voor een gemeten richting de bijdrage aan een kostenfunctie worden geformuleerd. Dat wordt hieronder in nader detail beschreven. Alvorens echter tot de (mathematische) details wordt overgegaan wordt nu eerst kort de achterliggende doelstelling vastgelegd. Die doelstelling is om bij een gegeven gemeten richting $\hat{\phi}$ een ISTM-vergelijking te definiëren die strikt in $\hat{\phi}$, V_{ϕ} (de onzekerheid van $\hat{\phi}$) en ϕ (de met ISTM gevonden richting, afgeleid uit de te berekenen transport vector $\vec{T} = (T_x, T_y)^T$) is uitgedrukt. Hierbij moeten voor de ISTM-gebruiker twee opties mogelijk zijn:

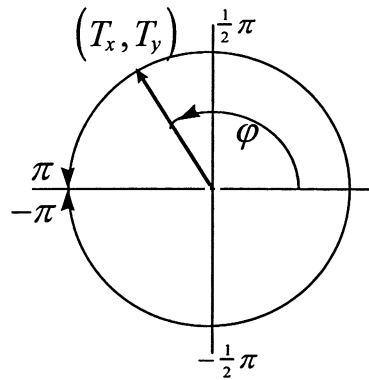
1. De door het ISTM te vinden richting ϕ moet zo goed mogelijk strikt parallel met de gemeten richting $\hat{\phi}$ zijn.
2. De door het ISTM te vinden richting ϕ moet zo goed mogelijk strikt parallel, òf zo goed mogelijk anti-parallel met de gemeten richting $\hat{\phi}$ zijn.

De consequenties voor het ISTM voor Optie 1 worden hieronder in Paragraaf 6.1 in volledig detail uitgewerkt. Op basis van het daar afgeleide formalisme wordt Optie 2 uitgewerkt in Paragraaf 6.2. Op dit moment (januari 2000) is Optie 1 daadwerkelijk geïmplementeerd in het niet lineaire ISTM, versie 2.00. Voor Optie 2 is dat echter nog niet het geval. Alvorens de technische achtergrond van de Opties 1 en 2 verder wordt uitgewerkt worden eerst een aantal aspecten t.a.v. transportrichtingen (zoals definities, notaties, eenheden, etc.) nader vastgelegd.

Tenzij anders vermeld, worden richtingen uitgedrukt in radialen. Een gemeten richting $\hat{\phi}$ of een met het ISTM berekende richting ϕ wordt geacht te zijn gespecificeerd volgens $-\pi < \hat{\phi} \leq \pi$, ofwel $\hat{\phi} \in (-\pi, \pi]$, en analoog $\phi \in (-\pi, \pi]$. Hierbij moet overigens worden opgemerkt dat het ISTM niet rekent met richtingen maar met transportvectoren. Zijn voor een of andere cel (waarvan we de spatiele indicering (m,n) hier weglaten) de ISTM-transporten gegeven door T_x en T_y (waarbij binnen de ISTM administratie $T_x = \Theta_k$ en $T_y = \Theta_{\ell}$ voor zekere k en ℓ) dan volgt de richting ϕ uit:

$$\phi = \text{atan2}(T_y, T_x) \quad (6.1)$$

Zie onderstaande figuur.



De functie $\text{atan2}(\cdot)$ is hierbij niet de inverse tangens functie (ofwel boogtangens, ofwel arctangens) $\text{bgtg}(\cdot)$ maar een intrinsieke functie zoals die in FORTRAN is gedefinieerd met de functie zoals met vergelijking (6.1) en bovenstaande figuur aangegeven. Deze $\text{atan2}(\cdot)$ is in feite een veralgemenisering van de $\text{bgtg}(\cdot)$ functie. Zo levert de $\text{atan2}(\cdot)$ functie uitkomsten $\varphi \in (-\pi, \pi]$, terwijl het bereik van de $\text{bgtg}(\cdot)$ functie het interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$

als bereik heeft. Afhankelijk van de positie van de vector $\vec{T} = (T_x, T_y)^T$ in een der vier kwadranten kan de formule van vergelijking (6.1) echter wel in een $\text{bgtg}(\cdot)$ formulering worden uitgedrukt. Die uitdrukkingen zullen in het vervolg expliciet benodigd zijn (namelijk bij het berekenen van afgeleiden van de hoek naar de componenten van de transportvektor) en worden derhalve hier nu expliciet weergegeven.

Stap 1 : Gegeven een transportvector $\vec{T} = (T_x, T_y)^T$ wordt de hoek ψ gedefinieerd door:

$$\begin{cases} T_x = 0, T_y > 0 & \rightarrow \psi = \frac{1}{2}\pi \\ T_x = 0, T_y < 0 & \rightarrow \psi = -\frac{1}{2}\pi \\ T_x \neq 0 & \rightarrow \psi = \text{bgtg}\left(\frac{T_y}{T_x}\right) \in (-\frac{1}{2}\pi, \frac{1}{2}\pi) \end{cases} \quad (6.2)$$

Stap 2a : Voor $\vec{T}$ in eerste en vierde kwadrant, d.w.z. als $T_x > 0$ dan $\varphi = \psi$

Stap 2b : Voor $\vec{T}$ in tweede kwadrant, d.w.z. als $T_x < 0, T_y > 0$ dan $\varphi = \psi + \pi$

Stap 2c : Voor $\vec{T} \neq \vec{0}$ in derde kwadrant, d.w.z. als $T_x < 0, T_y < 0$ dan $\varphi = \psi - \pi$

Als gevolg is voor elke $\vec{T} \neq \vec{0}$ de gezochte hoek φ van de vorm:

$$\varphi = \text{bgtg}\left(\frac{T_y}{T_x}\right) + k \cdot \pi \quad (6.3)$$

waarbij $k \in \{-1, 0, 1\}$. Binnen een kwadrant is deze k niet van $\vec{T}$ afhankelijk.

Met een gemeten richting $\hat{\varphi} \in (-\pi, \pi]$ en een ISTM richting $\varphi \in (-\pi, \pi]$ afgeleid volgens bovenstaand recept leidende tot vergelijking (6.3), zijn we nu voldoende bewapend om binnen het ISTM de conceptuele modelvergelijking voor richtingsinformatie te definiëren.

6.1 Optie 1. ISTM model voor strikt parallelle richtingen

Hoewel er diverse andere alternatieven mogelijk zijn wordt hier gekozen voor een zo eenvoudig mogelijke meetvergelijking t.a.v. hoeken, namelijk:

$$\varphi + V_\varphi = \hat{\varphi} \quad (6.1.1)$$

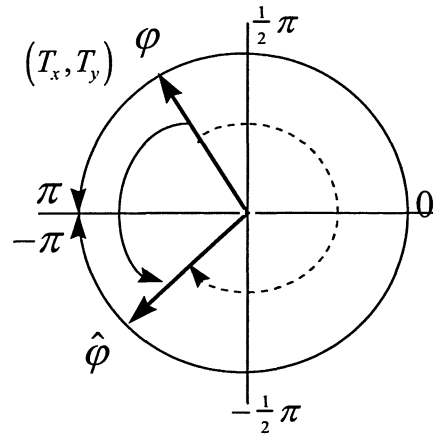
T.o.v. van eventuele alternatieve vormen heeft deze vergelijking namelijk het voordeel dat deze lineair is in de gemeten en te identificeren hoeken $\hat{\varphi}$ en φ en dus qua interpretatie zeer eenvoudig is. Bovendien is de onzekerheid V_φ in het rechterlid direct te ontleen aan de meetonzekerheid σ_φ van de hoek $\hat{\varphi}$ zonder dat die eerst vertaald hoeft te worden (hetgeen wel zou moeten indien het linkerlid een niet-lineaire expressie in φ is).

De interpretatie van deze vergelijking is dus dat de gemeten hoek $\hat{\varphi}$ de ‘echte’ hoek van de transport $\vec{T}$ vormt maar dan wel verstoord met een meetruis V_φ . Merk op dat deze meetvergelijking weliswaar lineair is in de hoeken, maar dat deze sterk niet-lineair is in de componenten van de transport vektor $\vec{T}$. Zie hiervoor vergelijking (6.3) van de vorige paragraaf die laat zien dat in dit verband een boogtangensfunctie voorkomt. Aangezien de ISTM-vergelijkingen zijn uitgedrukt in transporten, en naar deze grootheden moet worden opgelost, leidt vergelijking (6.1.1) dus tot een niet-lineaire (uitbreiding van het) ISTM.

Voor de oplosprocedure van niet-lineaire ISTM moet nu aan de hand van vergelijking (6.1.1) de bijdrage J_φ van een richtingsmeting $\hat{\varphi}$ aan de kostenfunctie worden gedefinieerd, alsmede diens gradiënt naar de te schatten $\vec{T}$ bij de (m,n) -positie van die $\hat{\varphi}$. Als de spreiding van die meetruis met σ_φ wordt aangegeven dan is (in eerste instantie) volgens de definitie van het niet-lineaire ISTM die bijdrage J_φ aan de totale kostenfunctie J_{ISTM} als volgt:

$$J_\varphi = \left(\frac{\varphi - \hat{\varphi}}{\sigma_\varphi} \right)^2 \quad (6.1.2)$$

Hierbij is nog steeds $\varphi, \hat{\varphi} \in (-\pi, \pi]$ en moet met name nog even goed gekeken worden naar de diskontinuiteit bij de overgang van het tweede naar het derde kwadrant. Als concreet voorbeeld die dit voorzichtig zijn rechtvaardigt laat $\varphi = \pi - \delta$ (in tweede kwadrant) en $\hat{\varphi} = -\pi + \varepsilon$ (in derde kwadrant) met δ, ε (kleine) positieve getallen. Dan liggen φ en $\hat{\varphi}$ in absolute zin dicht bij elkaar maar $\varphi - \hat{\varphi} = 2\pi - \delta - \varepsilon$ is relatief groot. Het onderliggende probleem is dat bij het vergelijken van twee hoeken we uitgaande van φ op twee manieren langs de eenheidscirkel naar $\hat{\varphi}$ kunnen wandelen, namelijk met een positieve draairichting (tegen de klok in) of met een negatieve draairichting (met de klok mee). Zie onderstaande figuur.



De moraal is dan de kortste weg te kiezen welke in absolute zin altijd kleiner of gelijk is π . Mathematisch gaat als volgt met behulp van de volgende definitie van $\Delta\varphi$:

$$\begin{cases} -\pi \leq \varphi - \hat{\varphi} \leq \pi & \rightarrow \Delta\varphi = \varphi - \hat{\varphi} \\ \varphi - \hat{\varphi} < -\pi & \rightarrow \Delta\varphi = \varphi - \hat{\varphi} + 2\pi \\ \varphi - \hat{\varphi} > \pi & \rightarrow \Delta\varphi = \varphi - \hat{\varphi} - 2\pi \end{cases} \quad (6.1.3)$$

Als gevolg is $\Delta\varphi$ van de vorm:

$$\Delta\varphi = \varphi - \hat{\varphi} + 2\ell \cdot \pi \in (-\pi, \pi] \quad (6.1.4)$$

Hierin is $\ell \in \{-1, 0, 1\}$. Als definitief resultaat, en met substitutie van de φ van vergelijking (6.3) wordt dan voor de bijdrage J_φ aan de totale kostenfunctie J_{ISTM} het volgende resultaat gevonden:

$$J_\varphi = \left(\frac{\Delta\varphi}{\sigma_\varphi} \right)^2 = \left(\frac{bgtg\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k + 2\ell) \cdot \pi}{\sigma_\varphi} \right)^2 \quad (6.1.5)$$

Hierbij kan worden opgemerkt dat in deze vorm J_φ maximaal is en ter waarde $\left(\frac{\pi}{\sigma_\varphi}\right)^2$ indien de φ en de $\hat{\varphi}$ anti-parallelle richtingen weergeven. Daarmee is een belangrijke tekortkoming van het lineaire ISTM verholpen.

Aangezien de afgeleide van de $bgtg()$ bepaald is volgens

$$f(x) = bgtg(x) \rightarrow f'(x) = \frac{1}{1+x^2} \quad (6.1.6)$$

volgt uit vergelijking (6.1.5) dat:

$$\begin{cases} \frac{\partial J_\varphi}{\partial T_x} = -2 \cdot \left(\frac{\text{bgtg}\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k+2\ell) \cdot \pi}{\sigma_\varphi^2} \right) \cdot \frac{T_y}{T_x^2 + T_y^2} = -2 \cdot \frac{\Delta\varphi}{\sigma_\varphi^2} \cdot \frac{T_y}{T_x^2 + T_y^2} \\ \frac{\partial J_\varphi}{\partial T_y} = 2 \cdot \left(\frac{\text{bgtg}\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k+2\ell) \cdot \pi}{\sigma_\varphi^2} \right) \cdot \frac{T_x}{T_x^2 + T_y^2} = 2 \cdot \frac{\Delta\varphi}{\sigma_\varphi^2} \cdot \frac{T_x}{T_x^2 + T_y^2} \end{cases} \quad (6.1.7)$$

Voor het bepalen van de bijdrage van de φ -vergelijkingen aan de (inverse) covariantie matrix P van een ISTM oplossing $\hat{\Theta}$ (zie de vergelijkingen 5.2.10-11 uit Hoofdstuk 5) moeten ook de tweede afgeleiden van de J_φ worden bepaald. Deze tweede afgeleiden vormen de zogenaamde Hessiaan van J_φ die geïnverteerd moet worden om de gewenste P te vinden. In Hoofdstuk 5 werd al gekonstateerd dat in de praktijk zo'n inversie niet mogelijk is gezien de omvang van die matrix. Derhalve wordt net zoals in het lineaire ISTM een benadering gemaakt door ons te beperken tot de diagonaal-elementen van de Hessiaan en deze dan te inverteren om de covariantiematrix P te vinden. Hiervoor moeten dan de $\frac{\partial^2 J_\varphi}{\partial T_x^2}$ en $\frac{\partial^2 J_\varphi}{\partial T_y^2}$ afgeleid worden. Uit (6.1.7) volgt dat:

$$\begin{cases} \frac{1}{2} \frac{\partial^2 J_\varphi}{\partial T_x^2} = \frac{1}{\sigma_\varphi^2} \cdot \frac{T_y}{(T_x^2 + T_y^2)^2} \cdot (T_y + 2 \cdot \Delta\varphi \cdot T_x) \\ \frac{1}{2} \frac{\partial^2 J_\varphi}{\partial T_y^2} = \frac{1}{\sigma_\varphi^2} \cdot \frac{T_x}{(T_x^2 + T_y^2)^2} \cdot (T_x - 2 \cdot \Delta\varphi \cdot T_y) \end{cases} \quad (6.1.8)$$

De consequenties van het niet meenemen van de $\frac{\partial^2 J_\varphi}{\partial T_x \partial T_y}$ is dat de varianties (ofwel de onzekerheden) van de optimale T_x en T_y onderschat worden, en er geen informatie beschikbaar is van de covarianties van de diverse T_x en T_y , d.w.z. de mate waarin deze grootheden van elkaar afhankelijk zijn.

Tenslotte worden nog kort de consequenties nagegaan voor een representatie van hoeken in graden in plaats van radialen. Hiertoe wordt de volgende faktor λ geïntroduceerd die radialen naar graden converteert:

$$\lambda = \frac{180}{\pi} \quad (6.1.9)$$

Als eerste stap wordt daarna de door een ISTM gebruiker in graden opgegeven richtingshoek $\hat{\varphi}$ (alsmede de onzekerheid σ_φ) gekonverteerd naar het bereik $(-180^\circ, 180^\circ]$ en daarna naar radialen door het eerste resultaat door de λ van vergelijking (6.1.9) te delen. Hierna wordt het hele formalisme gevolgd zoals hierboven is uitgewerkt voor radialen om de kostenfunctie J_φ en diens gradiënt J_φ te berekenen. Binnen deze representatie van hoeken in radialen vindt vervolgens ook de feitelijke minimalisatie van de

totale kostenfunctie J_{ISTM} plaats die in het bijzonder de diverse bijdragen J_φ van richtingsinformatie omvat. Op basis van de zo geïdentificeerde transporten $\vec{T}$ kunnen in de postprocessing weer richtingen worden afgeleid volgens de $\text{atan2}(\cdot)$ functie en conform vergelijking (6.1.9). De zo gevonden radialen kunnen in de uitvoer weer gemakkelijk naar graden geconverteerd worden.

Het is hierbij belangrijk om op te merken dat voor het uiteindelijke ISTM resultaat van transporten $\vec{T}$ het er niet toe doet welke representatie voor hoeken werd gekozen. Dat kan aan de kostenfunctie van vergelijking (6.1.5) worden herkend uit het feit dat het quotiënt $\frac{\Delta\varphi}{\sigma_\varphi}$ onafhankelijk is van de gehanteerde eenheid van hoeken. Derhalve dragen binnen de ISTM-integratie van informatiebronnen de richtingen altijd in dezelfde mate bij.

6.2 Optie 2. ISTM model voor parallelle of anti-parallelle richtingen

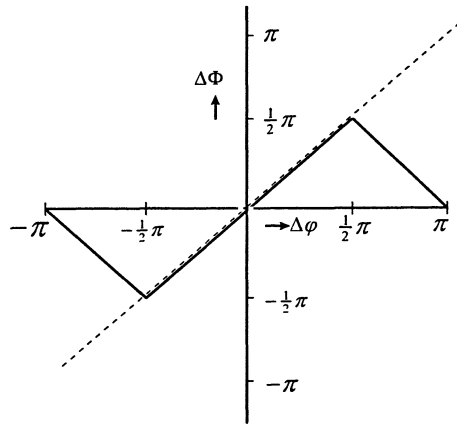
Bij deze optie wordt als uitgangspunt de voor Optie 1 afgeleide modelvergelijking (6.1.1) gehanteerd die kan worden geschreven als:

$$\Delta\varphi + V_\varphi = 0 \quad (6.2.1)$$

Hierin is $\Delta\varphi = \varphi - \hat{\varphi} + 2\ell \cdot \pi \in (-\pi, \pi]$ het verschil van de gemeten en berekende richtingen in radialen. Deze $\Delta\varphi$ is volgens vergelijking (6.1.4) in de vorige paragraaf afgeleid. Onder Optie 1 werd hieruit de niet-lineaire kostenfunctie van de vergelijkingen (6.1.2), respectievelijk (6.1.5) afgeleid. Deze kostenfunctie neemt kwadratisch toe wanneer $|\Delta\varphi|$ toeneemt van 0 (parallel) tot π (volledig anti-parallel). Voor Optie 2 kan dit verloop worden gehandhaafd zolang $|\Delta\varphi| \leq \frac{1}{2}\pi$. Voor $|\Delta\varphi| > \frac{1}{2}\pi$ moet daarentegen die kostenfunctie weer afnemen om bij 0 te eindigen wanneer $|\Delta\varphi|$ de waarde π bereikt. Dit kan worden bereikt door de $\Delta\varphi$ te corrigeren tot een $\Delta\Phi \in [-\frac{1}{2}\pi, \frac{1}{2}\pi]$ volgens het voorschrift:

$$\begin{cases} \Delta\Phi = -\pi - \Delta\varphi & \text{voor } -\pi < \Delta\varphi < -\frac{1}{2}\pi \\ \Delta\Phi = \Delta\varphi & \text{voor } -\frac{1}{2}\pi \leq \Delta\varphi \leq \frac{1}{2}\pi \\ \Delta\Phi = \pi - \Delta\varphi & \text{voor } \frac{1}{2}\pi < \Delta\varphi \leq \pi \end{cases} \quad (6.2.2)$$

De doorgetrokken grafiek in onderstaande figuur geeft een illustratie van $\Delta\Phi$ als functie van $\Delta\varphi$. De gestippelde grafiek geeft de identieke functie en is zoals gebruikt onder Optie 1.



Voor Optie 2 is dan de richtingsvergelijking voor het conceptuele model als volgt:

$$\Delta\Phi + V_\varphi = 0 \quad (6.2.3)$$

De bijdrage van deze modelvergelijking aan de kostenfunctie is dan:

$$J_\varphi = \left(\frac{\Delta\Phi}{\sigma_\varphi} \right)^2 = \begin{cases} \left(\frac{-\pi - \Delta\varphi}{\sigma_\varphi} \right)^2 & \text{voor } -\pi < \Delta\varphi < -\frac{1}{2}\pi \\ \left(\frac{\Delta\varphi}{\sigma_\varphi} \right)^2 & \text{voor } -\frac{1}{2}\pi \leq \Delta\varphi \leq \frac{1}{2}\pi \\ \left(\frac{\pi - \Delta\varphi}{\sigma_\varphi} \right)^2 & \text{voor } \frac{1}{2}\pi < \Delta\varphi \leq \pi \end{cases} \quad (6.2.4)$$

Ofwel:

$$J_\varphi = \begin{cases} \left(\frac{\Delta\varphi + \pi}{\sigma_\varphi} \right)^2 & \text{voor } -\pi < \Delta\varphi < -\frac{1}{2}\pi \\ \left(\frac{\Delta\varphi}{\sigma_\varphi} \right)^2 & \text{voor } -\frac{1}{2}\pi \leq \Delta\varphi \leq \frac{1}{2}\pi \\ \left(\frac{\Delta\varphi - \pi}{\sigma_\varphi} \right)^2 & \text{voor } \frac{1}{2}\pi < \Delta\varphi \leq \pi \end{cases} \quad (6.2.5)$$

Met

$$m := \begin{cases} 1 & \text{voor } -\pi < \Delta\varphi < -\frac{1}{2}\pi \\ 0 & \text{voor } -\frac{1}{2}\pi \leq \Delta\varphi \leq \frac{1}{2}\pi \\ -1 & \text{voor } \frac{1}{2}\pi < \Delta\varphi \leq \pi \end{cases} \quad (6.2.6)$$

is dan:

$$J_{\varphi} = \left(\frac{\Delta\varphi + m \cdot \pi}{\sigma_{\varphi}} \right)^2 \quad (6.2.7)$$

Uitgedrukt in transporten, gebruikmakende van de in de voorafgaande paragrafen afgeleide definities voor de integers k en ℓ wordt dit dan tenslotte:

$$J_{\varphi} = \left(\frac{bgtg\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k + 2\ell + m) \cdot \pi}{\sigma_{\varphi}} \right)^2 \quad (6.2.8)$$

Voor $m=0$ komt deze vorm overeen met vergelijking (6.1.5) van Paragraaf 3.1. De eerste en tweede orde afgeleides van J_{φ} zoals die onder de vergelijkingen (6.1.7) en (6.1.8) zijn afgeleid vertalen zich voor het huidige geval dan naar:

$$\begin{cases} \frac{\partial J_{\varphi}}{\partial T_x} = -2 \cdot \left(\frac{bgtg\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k + 2\ell + m) \cdot \pi}{\sigma_{\varphi}^2} \right) \cdot \frac{T_y}{T_x^2 + T_y^2} = -2 \cdot \frac{\Delta\varphi + m \cdot \pi}{\sigma_{\varphi}^2} \cdot \frac{T_y}{T_x^2 + T_y^2} \\ \frac{\partial J_{\varphi}}{\partial T_y} = 2 \cdot \left(\frac{bgtg\left(\frac{T_y}{T_x}\right) - \hat{\varphi} + (k + 2\ell + m) \cdot \pi}{\sigma_{\varphi}^2} \right) \cdot \frac{T_x}{T_x^2 + T_y^2} = 2 \cdot \frac{\Delta\varphi + m \cdot \pi}{\sigma_{\varphi}^2} \cdot \frac{T_x}{T_x^2 + T_y^2} \end{cases} \quad (6.2.9)$$

en

$$\begin{cases} \frac{1}{2} \frac{\partial^2 J_{\varphi}}{\partial T_x^2} = \frac{1}{\sigma_{\varphi}^2} \cdot \frac{T_y}{(T_x^2 + T_y^2)^2} \cdot (T_y + 2 \cdot (\Delta\varphi + m \cdot \pi) \cdot T_x) \\ \frac{1}{2} \frac{\partial^2 J_{\varphi}}{\partial T_y^2} = \frac{1}{\sigma_{\varphi}^2} \cdot \frac{T_x}{(T_x^2 + T_y^2)^2} \cdot (T_x - 2 \cdot (\Delta\varphi + m \cdot \pi) \cdot T_y) \end{cases} \quad (6.2.10)$$

7 Discussie en aanbevelingen

Geïnitieerd door de ervaringen die door RIKZ zijn opgedaan in diverse, grootschalige praktijktoepassingen met het ISTM (versie 1.03) is in het hier voorliggende onderzoek in nader detail geïnventariseerd of het modelconcept op de volgende drie punten zou kunnen worden verbeterd:

1. Het inbrengen van meer fysika in het model
2. De (her)definitie van de wijze waarop transportrichtingen in het model verwerkt worden
3. De functie van regularisatie in het ISTM

Als algehele conclusie van deze studie kan worden gesteld dat bij al deze aspecten diverse mogelijkheden voor de gewenste verbeteringen worden voorzien.

Voor het tweede aandachtspunt heeft dit geresulteerd in een uitvoerige beschrijving van de noodzakelijke aanpassingen in het conceptuele model (zie Hoofdstuk 5 en 6). Deze aanpassingen zijn tevens geïmplementeerd en aan de hand van testen met positief resultaat gevalideerd. Dit heeft geleid tot de operationalisatie van versie 2.0 van het ISTM. Een belangrijk verschil met de voorafgaande versie 1.03 is dat het model met niet-lineaire vergelijkingen voor het verwerken van richtingsinformatie is uitgebreid geworden. Per gemeten richting kunnen twee opties worden opgegeven namelijk of er wél of géén onderscheid moet worden gemaakt tussen hoeken met een verschil van 180° . Overigens kan Versie 2.0 zodanig worden aangestuurd dat deze uitbreidingen niet geactiveerd worden en het systeem functioneel reduceert tot dat van Versie 1.03.

Voor de overige twee aandachtspunten zijn de resultaten na een uitvoerige discussie kort en puntsgewijs samengevat in lijsten met aanbevelingen. Zie hiervoor Paragraaf 2.4 op bladzijde 2-6 en Paragraaf 4.4 op bladzijde 4-17. Op de daar ingebrachte aandachtspunten en hun prioriteitenstelling zal in dit hoofdstuk niet meer op individuele basis worden teruggegrepen. Veelmeer zullen hier een aantal aanbevelingen worden geformuleerd die daarvan een meer algemene synthese vormen.

Op basis van een nadere evaluatie van de overeenkomsten en verschillen van de diverse individuele aanbevelingen ziet die synthese er dan puntsgewijs als volgt uit.

1. Voor zover de diverse vanuit *fysische voorkennis* en *regularisatie* ingebrachte aspecten voor verbetering met *lineaire vergelijkingen in de transportcomponenten* zijn weergegeven, kunnen deze met één generieke uitbreiding van het ISTM gerealiseerd worden. Elke vergelijking heeft namelijk de volgende algemene vorm

$$\sum_{\ell=1}^L \alpha_{\ell} \cdot \Theta_{k_{\ell}} + V = R \quad (7.1)$$

Elk der $\Theta_{k_{\ell}}$ is een der componenten $(T_x)_m^n$ of $(T_y)_m^n$ voor een bepaalde cel (m,n) van het numerieke rooster. L is het aantal transportcomponenten dat met zo'n vergelijking wordt gekoppeld, en de α_{ℓ} geven de 'stempel' van die koppelingen. R is het bijbehorende rechterlid van de vergelijking, en de ruis V geeft de onzekerheid van dat rechterlid (en moet aan de hand van een spreiding σ gekwantificeerd worden). Voor elke (fysische of regularisatie) modelvergelijking zullen het aantal en de waarde van deze grootheden variëren.

Met de algemene vergelijking (7.1) zijn dan niet alleen de oude vormen van meet- en regularisatievergelijkingen omvat maar ook de in de vorige hoofdstukken van dit rapport gegeven nieuwe vormen, ongeacht of deze alleen lokale/naburige transporten koppelen (zoals de continuïteitsvergelijkingen en 1^e en 2^e orde regularisatievergelijkingen die aan individuele rekencellen zijn gekoppeld), of meer ruimtelijk gedistribueerd van aard zijn (zoals vergelijkingen die aggregaties van transporten of bergingen bevatten).

In het bijzonder kunnen op deze manier door de gebruiker vrijwel willekeurige vormen van regularisaties worden opgegeven. Deze kunnen t.a.v. hun gewicht (ofwel de spreiding) ruimtelijk gevarieerd worden.

Omdat op deze manier al heel veel van de voorgestelde ISTM-verbeteringen gerealiseerd kan worden wordt aanbevolen deze generieke uitbreiding aan de ISTM-implementatie toe te voegen. Hierbij moeten twee kanttekeningen worden geplaatst. De eerste hiervan is dat in nader detail nagegaan zal moeten worden hoe deze uitbreidingen informatica-technisch georganiseerd moeten worden, vooral met het oog op de invoer-flexibiliteit waarmee de gebruiker deze kan specificeren. De tweede kanttekening betreft de consequenties voor het ISTM-oplosmechanisme; een eerste evaluatie lijkt echter aan te geven dat deze beperkt zijn. Beide punten pleiten in ieder geval voor het in eerste instantie opstellen van een voldoende gedetailleerd technisch ontwerp alvorens aan de voorgestelde implementatie wordt begonnen.

2. Afgezien van de niet-lineaire uitbreiding van het ISTM die werd geïnitieerd door richtingsvergelijkingen, lijken er op basis van de overige bevindingen van deze voorstudie vooralsnog geen voldoende argumenten aanwezig om andere *niet-lineaire* uitbreidingen van het ISTM te ontwerpen en te implementeren.

Voor een conform het bovenstaande uitgebreid ISTM wordt t.a.v. diens gebruik in praktijktoepassingen nog het volgende opgemerkt en/of aanbevolen.

3. In de praktijk zijn ISTM toepassingen dermate gekompliceerd dat het niet duidelijk is of onder dergelijke condities de in dit onderzoek voorgedragen ideeën voor de verbetering van het ISTM daadwerkelijk perspectieven bieden. Daarom wordt aanbevolen om dit onder gekontroleerde omstandigheden in meer expliciete en kwantitatieve vorm te toetsen. Hierbij wordt gedacht aan ISTM toepassingen op data (transporten en bergingen) afkomstig van een morfologisch model waarmee zand of sediment balansen van één of meer complexe (bestaande of fiktieve) prototype systemen zijn gesimuleerd. Bij het toetsen van de performance van het ISTM vormen de uitkomsten van het morfologische model dan een geschikte referentie. Behalve het feitelijke toetsen van de hier voorgestelde verbeteringen (t.a.v. inbrengen van meer fysische kennis en alternatieve regularisatiemechanismen) kan met deze ‘twin’ experimenten belangrijke ervaring worden opgedaan voor het ISTM t.a.v. minimale data behoeften, identificeerbaarheid van lokale en/of geaggregeerde transportpatronen, éénduidigheid van de gevonden transportvelden, gevoeligheden voor randvoorwaarden en lokale of gedistribueerde verstoringen, te kiezen vorm en sterkte van regularisaties, etc. De ervaringen die in deze twin experimenten wordt opgedaan is dan ‘later’ in de praktijk direct over te hevelen naar ‘echte’ case studies, zowel *a priori* (d.w.z. bij het vaststellen van de haalbaarheid van zo’n case, en de te plegen aansturingen van het ISTM gegeven de doelstelling en beschikbare meetgegevens en gebieds/systeem-kennis), als *a posteriori* (interpretatie en analyse van de ISTM-uitvoer, alsmede de feitelijke waarde die daaraan toegekend kan worden).
4. Aan de voorafgaande aandachtspunten kan worden toegevoegd dat het bij praktijktoepassingen niet altijd mogelijk zal zijn om *kwantitatieve* transportinformatie te vergaren uit fysische-, data-georiënteerde- en/of gebiedskennis. Vaak zal men niet verder komen dan indicaties of ruwe schattingen voor grootheden die als additionele

informatie/invoer aan het ISTM-kunnen worden aangeboden. Door hierin geschikte variaties aan te brengen kan echter inzicht worden verkregen in de lokale en globale gevoeligheden van de ISTM-resultaten, alsmede (in)consistenties met de overige (basis)gegevens. Uit de in zo'n gevoeligheidsanalyse gegenereerde serie transportvelden kan door een expert eventueel alsnog een meest realistische variant geselecteerd worden.

5. Een van de aspecten die veelvoudig in deze studie naar het verbeteren van het ISTM naar voren is gekomen is het toevoegen van geaggregeerde informatie (van transporten, richtingen, of bergingen) aan de invoer van het model. Omgekeerd zouden dan ook de resultaten van een ISTM berekening op dat niveau geïnterpreteerd en getoetst moeten (kunnen) worden. Op dit moment is die uitvoer in volledig gedistribueerde vorm waarbij per cel de transportvektor en berging wordt gegenereerd. Hieruit kunnen echter relatief eenvoudig de gewenste geaggregeerde grootheden worden afgeleid. Die zouden bijvoorbeeld kunnen bestaan uit balansen op het niveau van deelgebieden/morfologische eenheden. De concrete aanbeveling is dan om geschikte postprocessings-tools te ontwikkelen die de gebruiker toegang geven tot de gewenste geaggregeerde grootheden en op deze manier helpen voor een meer grootschalige interpretatie/presentatie/validatie/sturing/etc. van de ISTM-resultaten.

Op basis van de in Appendix B beschreven testtoepassingen van het niet-lineaire richtingen-formalisme zoals dat konform de beschrijving van Optie 1 in het ISTM (versie 2.00) is geïmplementeerd kan het volgende nog worden toegevoegd.

6. Zoals mocht worden verwacht houdt het niet-lineaire ISTM veel dwingender rekening met richtingsinformatie dan het lineaire ISTM. Verder bleek in deze case dat het uiteindelijk gevonden transportenveld in sterke matig gevoelig was voor de gekozen onzekerheden van de diverse invoergrootheden en/of de initial guess van het transportenveld. Deze grote gevoeligheid en niet éénzijdigheid van de oplossing is echter het directe resultaat van de gebruikte invoer data set (transportgroottes voorgeschreven aan een linker en rechterrاند, en alleen transportrichtingen in één centrale raai), en niet zozeer een vorm van falen van het inverse model. Van te voren kon namelijk al op analytische basis het bestaan van meerdere oplossingen worden aangetoond, en het inverse probleem is dus in feite 'ill posed'. Indien in toekomstige praktijktoepassingen vergelijkbare ervaringen worden opgedaan (waarbij de modelgeometrie en de spatiële verdeling van de invoergegevens veel complexer zijn, en niet meer *a priori* op analytische basis een uitspraak kan worden gedaan over het wel of niet goed gesteld zijn van het inverse probleem) kan hieraan een diagnose of hypothese worden ontleend dat de ISTM invoer in kwalitatieve en/of kwantitatieve zin niet toereikend is. Op dat moment kan of moet, al dan niet konform de diverse in dit rapport geformuleerde aanbevelingen, alsnog additionele (fysische) voorkennis worden aangedragen. Eens te meer laat de testtoepassing daarmee zien dat het ISTM zeker niet als black box moet worden gebruikt en dat telkens de nodige voorzichtigheid moet worden betracht bij het opstellen van de doelstelling van een case, diens haalbaarheid, en de inrichting van het daaruit volgende traject van het selekteren en samenstellen van de invoergegevens tot aan een kritische analyse en interpretatie van de uitvoer-resultaten.

Referenties

- Alkyon & WL | Delft Hydraulics, 1999, "Morphodynamic simulations of the tidal inlet "Zeegat van Ameland; Report 1: Validation of Delft2D-MOR (in English); Rapport 2: modelberekeningen (in Dutch)", Report A505/Z2731.
- Allersma, E., 1988, "Morfologisch onderzoek Noordelijk Deltabekken", Waterloopkundig Laboratorium | WL, Rapport Z71-03 (bestaat uit vier delen).
- De Vriend, H.J., 1997. "Prediction of aggregated-scale coastal evolution", in: Coastal Dynamics '97, ASCE, New York (in press).
- De Vriend, H.J., Capobianco, M., Chesher, T., De Swart, H.E., Latteux, B. and Stive, M.J.F., 1993, "Approaches to long-term modelling of coastal morphology: a review", Coastal Engineering, Vol 21, No 1-3, pp 225-269.
- De Vriend, H.J., 1998, "Prediction of aggregated-scale coastal evolution, (PACE)", Third European Marine Science and Technology Conference
- Eysink, W.D., 1990, "Morphologic response of tidal basins to changes", Proc. 22nd Int. Congress on Coastal Engineering, Chapter 147, Paper No 8, pp 1948-1961.
- Eysink, W.D., 1992, "Impact of sea level rise on the morphology of the Wadden Sea in the scope of its ecological function", Delft Hydraulics, Report H1300, phase 3.
- Jeuken, M.J.L., 2000, "Morphodynamic behaviour of tidal channels in the Western Scheldt", PhD-thesis, Delft University of Technology (in preparation).
- Koster, A.W.J. and Van den Boogaard, H.F.P., 1993a, "Invers Sedimenttransportmodel. Definitiestudie", Delft Hydraulics, Report Z632.10
- Koster, A.W.J. and Van den Boogaard, H.F.P., 1993b, "Invers Sedimenttransportmodel. Ontwerp", Delft Hydraulics, Report Z632.20
- Latteux, B., 1995, "Techniques for long-term morphological simulation under tidal current action", Marine Geology, Vol 126, pp 129-141.
- Marion, B.B. van, 1999, "Zandbalans van het Zeegat van Texel met het Invers Sediment Transport Model (1931 tot 1997)", Deel I en Deel II: Appendix. RIKZ/OS-99.116x
- Mulder, H.P.J., 1996, "Sedimentbalans Eems-Dollard: een tweedimensionale rekenmethode. Een zand- en slibbalans voor de periode 1985-1990 met het Invers Sediment Transport Model.". Rapport RIKZ-96.013.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1994, "Numerical Recipes in FORTRAN. The Art of Scientific Computing. Second Edition". Cambridge University Press, Cambridge.
- Roelvink, J.A. and Wang, Z.B., 1996, "Morphological Model System - Evaluation of DELFT2D-MOR", WL | Delft Hydraulics, Report Z2064.
- Roelvink, J.A., 1999, "Validation of Delft2D-MOR", WL | Delft Hydraulics, report Z2731 (draft).
- Siefert, W., 1987a, "Umsatz- und Bilanz-Analysen für das Küstenvorfeld der Deutschen Bucht - Grundlagen und erste Auswertungen (Teil I der Ergebnisse eines KFKI-Projektes)", in "Die Küste", Archiv für Forschung und Technik an der Nord- und Ostsee, Sonderdruck aus Heft 45/1987, pp 1-57.
- Siefert, W. en Lassen, H., 1987b, "Zum säkularen Verhalten der mittleren Wathöhen an ausgewählten Beispielen", in "Die Küste", Archiv für Forschung und Technik an der Nord- und Ostsee, Sonderdruck aus Heft 45/1987, pp 59-70.
- Stive, M.J.F., Wang, Z.B., Capobianco, M., Ruol, P. and Buijsman, M.C., 1998, "Morphodynamics of a tidal lagoon and the adjacent coast", in: Physics of Estuaries and Coastal Seas, ed. by J. Dronkers & M. Scheffers, Balkema, Rotterdam, pp. 397-407.
- Thoolen, P.M.C. and Wang, Z.B., 1999, "Sedimenttransportmodellering Westerschelde", WL | Delft Hydraulics, Report Z2489.
- Van den Boogaard, H.F.P., 1994, "Invers Sedimenttransportmodel. Toepassingen", Delft Hydraulics, Report Z632.
- Van den Boogaard, H.F.P. and Mulder, H.P.J., 1997, "Inverse Modelling of Sedimenttransport". In: Environmental and Coastal Hydraulics: Protecting the Aquatic Habitat, Volume 2 Proceedings of Theme B, Water for A Changing Global Community, The 27th Congress of the International Association for Hydraulic Research, San Francisco, August 10-15, 1997, ed. by Wang and Carstens, pp 1191-1196.
- Veen, J. van, 1950, "Eb- en vloed-schaarsystemen in de Nederlandse getijwateren", KNAG Tijdschrift, 2e reeks, deel 67, pp 303-325.

Wang, Z.B., Karssen, B., Fokkink, R.J. and Langerak, A., 1998, "A dynamic-empirical model for estuarine morphology", in: *Physics of Estuaries and Coastal Seas*, ed. by J. Dronkers & M. Scheffers, Balkema, Rotterdam, pp. 279-286.

WL | Delft Hydraulics, 1996, "Manual Delft2d-mor, release 2.07, an introduction to DELFT2D-MOR", november 1996.

APPENDICES

A User manual ISTM, Versie 2.00

A.1 Wijzigingen in de invoerbestanden

In Paragraaf A.2 van dit appendix is een voorbeeld te vinden van een nieuw invoerbestand voor het ISTM, versie V2.00. De wijzigingen t.o.v. de vorige versie, ISTM V1.03, bevinden zich met name aan het einde van dit bestand, en omvatten:

1. Aangepast invoerblok voor de lineaire solver
2. Extra invoerblok voor de niet-lineaire solver
3. Extra invoerblok voor richtingsspreiding
4. Extra invoerblok voor begintoestand van model

Conform bovenstaande volgorde worden hieronder de gewijzigde/toegevoegde invoerblokken in nader detail besproken. De gearceerde gedeelten uit de tekst zijn verbatim overgenomen uit een ISTM invoerbestand. Waar nodig wordt het gebruik van de invoerparameters in meer uitgebreide vorm toegelicht.

Ad 1: Invoerblok voor lineaire solver:

* Lineaire Solvers [SLAP]			
* Solver	Aantal iteraties	Nauwkeurigheid	
* ISolver	MaxIter	Eps	
1	15000	1.E-3	

ISolver:

Niet-negatieve integer voor de keuze van de solver voor het lineaire systeem ("Lineaire Solver")

- =0 : Solver staat uit; dus lineaire systeem wordt niet opgelost
ISTM begint dan direkt met het oplossen van het niet-lineaire systeem
- =1 : Conjugeerde Gradiënten Methode (idem V1.03)
- =2 : Bi-Conjugeerde Gradiënten Methode (idem V1.03)
- =3 : GMRES-Methode (idem V1.03)
(Generalised Minimal Residual Methode)

Advieswaarde: ISolver = 1

MaxIter:

Integer met maximaal aantal iteraties voor de lineaire solver. Geadviseerd wordt dit maximum op enkele duizenden iteraties te zetten. Feitelijk fungeert dit niet als het normale stopcriterium: eigenlijk dient het iteratie-proces te stoppen op het nauwkeurigheds-criterium (eps). Het kan echter gebeuren dat daarvoor dan een zo groot aantal iteraties benodigd is dat de rekentijd onaanvaardbaar lang wordt. Zie ook onderstaande toelichting.

Eps:

Positieve real voor de nauwkeurigheid waarmee het minimum wordt bepaald. Wanneer de gradient-norm $< \text{Eps}$ wordt het iteratie-proces gestopt. Voor de norm van de gradient wordt in de SLAP solver de maximum-norm gehanteerd. Advieswaarde voor Eps ligt in de range 1.E-3 - 1.E-6

Toelichting:

Ook in ISTM V2.00 is de lineaire solver nog steeds van belang, maar nu om een goede startwaarde te vinden voor het oplossen van het niet-lineaire systeem. Het wordt sterk afgeraden om direkt vanaf de beginsituatie de niet-lineaire solver (zie verder) aan te zetten, omdat er hierdoor minder garantie bestaat op het vinden van het minimum. Bovendien is ook de lineaire solver (per iteratie-stap) een stuk sneller. Feitelijk moet de lineaire solver worden gehanteerd om een goede 'warme start' voor de niet-lineaire solver te verkrijgen.

Om de voortgang van het iteratie-proces te kunnen volgen, wordt de foutparameter E op het scherm (en printbestand) getoond. Het normale gedrag is dat deze foutparameter geleidelijk (soms echter oscillerend) afneemt, totdat deze kleiner is dan het opgegeven foutcriterium EPS, en het iteratie-proces wordt beëindigd. Op deze manier geeft dit een goed beeld over de voortgang van het iteratie-proces. (*niet nieuw: dit was al zo in ISTM V1.03*)

Voorbeeld van de scherm-uitvoer bij gebruik van de SLAP-solver:

```
Iter= 7641 E= 0.45E-04 [Mx= 0.10E-04] I      *
Iter= 7651 E= 0.89E-04 [Mx= 0.10E-04] I      *
Iter= 7661 E= 0.11E-03 [Mx= 0.10E-04] I      *
Iter= 7671 E= 0.48E-04 [Mx= 0.10E-04] I      *
Iter= 7681 E= 0.54E-04 [Mx= 0.10E-04] I      *
Iter= 7691 E= 0.46E-04 [Mx= 0.10E-04] I      *
```

De print-plot aan het eind geeft de foutparameter E weer. Wanneer deze de I-as raakt, is dat het einde van het iteratie-proces.

Ad 2: Invoerblok voor niet-lineaire solver:

* Niet Lineaire Solvers			
*	Solver	Aantal iteraties	Nauwkeurigheid
*	ISolver	MaxIter	Eps
	1	15000	1.E-4

ISolver:

Positieve integer voor de keuze van de solver voor het niet-lineaire systeem ("Niet-lineaire Solver")

- =0 : Solver staat uit; dus niet-lineaire systeem wordt niet opgelost
Te gebruiken wanneer men alleen in de resultaten van het lineaire systeem geïnteresseerd is.
- =1 : Quasi Newton (Limited memory versie)
Deze methode blijkt (nog) niet voldoende robuust te functioneren voor grote systemen; geadviseerd wordt deze optie niet te gebruiken.
- =2 : Geconjugeerde Gradiënten Methode (voor niet-lineaire stelsels)

Advieswaarde: ISolver = 2

Opmerking. Zijn zowel de lineaire en de niet-lineaire solver geactiveerd in het invoerbestand dan wordt nadat de lineaire solver klaar is de niet-lineaire solver automatisch gestart. Deze niet-lineaire solver gebruikt dan het met de lineaire solver gegenereerde transporten - en bergingenveld als initial guess.

MaxIter:

Positieve integer voor het maximale aantal iteraties voor de niet-lineaire solver. Geadviseerd wordt het maximum op minstens enkele duizenden iteraties te zetten. Feitelijk is dit niet het eigenlijke stopcriterium: normaliter dient het iteratie-proces te stoppen op het nauwkeurigheds-criterium (eps). Zie ook de overige toelichtingen hieronder.

Eps:

Positieve real voor de nauwkeurigheid van het te vinden minimum. Wanneer de gradiënt-norm² < eps wordt het iteratie-proces gestopt. Voor de gradiënt-norm wordt hier de (Euclidische) L₂-norm genomen. Advieswaarde: range 1.E-3 - 1.E-6

Toelichting:

Vanuit de formulering van het niet-lineaire ISTM systeem volgt dat de interpretatie van (de voortgang van) het iteratie-proces voor de niet-lineaire solvers een stuk lastiger is dan voor de lineaire solvers. Om toch enig inzicht te hebben in de voortgang, worden zowel de kostenfunctie als de norm van de gradiënt op het scherm getoond.

Vanuit de aard van het minimalisatie-proces moet de kostenfunctie een strikt monotoon-afnemend gedrag vertonen. Voor de gradiënt-norm mag eveneens worden verwacht dat deze, overall gezien, afneemt (voor het minimum geldt immers dat de gradiëntvektor identiek aan de nulvektor moet worden). Dat gedrag hoeft echter niet strikt monotoon-afnemend te zijn en kan ook een oscillerend-afnemend verloop hebben.

In toepassingen (o.a. estuarium-model) is gebleken, dat de grootte-orde van kostenfuncties en gradiënten voor het lineaire en het niet-lineaire deel nogal sterk uiteen kunnen lopen. Wanneer je deze op het scherm als totale kostenfunctie (lineair + niet-lineair) presenteert, dan kan het gebeuren dat het tijdens het minimaliseren lijkt 'of er niets gebeurt'. In werkelijkheid is de solver dan bezig om de ondergeschikte kostenfunctie (lineair of niet-lineair) te minimaliseren, met slechts zeer geringe reductie van de totale kostenfunctie!

Vanuit deze overwegingen is ervoor gekozen zowel de toestand van het lineaire als van het niet-lineaire deel van het systeem zichtbaar te maken in de vorm van onderstaande kostenfuncties en de gradiënten:

- F_l kostenfunctie voor het lineaire deel
- F_{nl} kostenfunctie voor het niet-lineaire deel
- G_l gradiënt-norm voor het lineaire deel
- G_{nl} gradiënt-norm voor het niet-lineaire deel

Zoals bekend, staat in ISTM V2.00 het niet-lineaire deel voor de verschillen van gemeten en berekende richtingen. De kostenfunctie F_{nl} is dan de som van de (gewogen) kwadratische verschillen, ofwel $F_{nl} = \sum_{k=1}^K \left(\frac{\Delta \varphi_k}{\sigma_k} \right)^2$. De overige vergelijkingen zijn lineair in de transporten en worden aan de kostenfunctie F_l voor het lineaire deel toegevoegd.

In toepassingen blijkt dat het (helaas) meestal niet lukt om een dusdanige afschatting van de parameter EPS te vinden, dat het minimalisatie-proces ook daadwerkelijk op een 'nette' manier wordt gestopt: d.w.z. op het nauwkeurigheds-criterium.

De volgende aanpak lijkt (voorlopig) het beste:

- kies voor EPS een 'kleine' waarde conform SLAP solver
- kies voor het maximaal aantal iteraties enkele (tien-)duizenden iteraties

Gedurende het minimalisatie-proces probeert de solver langs een berekende zoeklijn (bepaald volgens geconjugeerde gradiënten) een nog lager minimum te vinden. Op basis van de bovenstaande aanpak, zal het effect zijn dat de niet-lineaire solver stopt in een situatie waarin het langs de berekende zoeklijn niet lukt om een nog kleinere waarde te vinden. Dit betekent dat uiteindelijk het minimum van de kostenfunctie is gevonden. De solver geeft hiervoor de melding:

WARNING ERROR 1 from DU2CGG the line search of an integration was abandoned

Voorlopig moet dit worden beschouwd als een normaal stopcriterium, d.w.z. dat het minimum is gevonden. Het blijft natuurlijk altijd van belang een primaire check uit te voeren in de uitvoerbestanden, of dit resultaat ook plausibel is. Een eerste controle zou kunnen bestaan uit het nagaan van de individuele verschillen tussen meting en berekening, zoals die te vinden zijn in het STAT uitvoerbestand. Verder is het van belang de kostenfuncties te controleren.

Er moet derhalve niet zonder meer vanuit worden gegaan dat ook daadwerkelijk het 'echte' minimum gevonden is. Veel meer nog dan voor de SLAP solvers, blijft het controleren van de oplossing van groot belang!

Voorbeeld scherm-uitvoer Conjugeerde Gradiënt (CGM) solver:

```
6 Fl=0.117E+04 Fnl=0.602E-06 |Gl|max=0.500E+05 |Gnl|max=0.233E+05
7 Fl=0.117E+04 Fnl=0.602E-06 |Gl|max=0.500E+05 |Gnl|max=0.336E+06
8 Fl=0.117E+04 Fnl=0.379E-03 |Gl|max=0.196E+05 |Gnl|max=0.767E+10
9 Fl=0.117E+04 Fnl=0.379E-03 |Gl|max=0.196E+05 |Gnl|max=0.905E+09
10 Fl=0.117E+04 Fnl=0.384E-03 |Gl|max=0.198E+05 |Gnl|max=0.131E+13
11 Fl=0.117E+04 Fnl=0.384E-03 |Gl|max=0.198E+05 |Gnl|max=0.117E+13
12 Fl=0.117E+04 Fnl=0.384E-03 |Gl|max=0.198E+05 |Gnl|max=0.238E+13
13 Fl=0.117E+04 Fnl=0.384E-03 |Gl|max=0.198E+05 |Gnl|max=0.843E+13
```

Ad 3: Invoerblok voor niet-lineaire richtingspreiding

```
* Niet-lineaire richtingspreiding
* Sigma
1.0
```

Sigma:

Positieve real voor uniforme richtingspreiding (in graden) voor het niet-lineaire systeem.

Toelichting:

Bij het oplossen van het lineaire systeem worden de spreidingen toegepast zoals die op de meetbestanden staan voor de transporten. Voor het oplossen van het niet-lineaire systeem wordt de in dit blok opgegeven uniforme spreiding gebruikt: alle gemeten richtingen hebben dezelfde spreiding.

Ad 4: Invoerblok voor startwaarden

* Startwaarden voor Tx en Ty
* Dominante transportgrootte : 100 ton/s
* Dominante transportrichting: oost
* Tx (ton/s) Ty (ton/s)
100 0

T_x :

Uniforme startwaarde voor transportkomponent T_x .

T_y :

Uniforme startwaarde voor transportkomponent T_y .

Toelichting:

In ISTM V1.03 vormde de rusttoestand van het model ($\vec{T}(\vec{r}) = \vec{0}$ voor alle $\vec{r}$) de beginsituatie voor het oplossen van het (lineaire) systeem. Software-matig stonden hierbij de startwaarden voor transportcomponenten T_x en T_y beide op 0.0.

Er zijn belangrijke redenen waarom het van belang is om als gebruiker de begintoestand te kunnen opgeven.

- De keuze van de begintoestand voor de transporten is van belang, om zoveel mogelijk informatie te kunnen meegeven aan het ISTM model, zowel wat betreft transportgrootte als qua richting. Naarmate de begintoestand meer overeenkomt met de aangeleverde metingen, mag verwacht worden dat de solver beter (robuuster) en sneller in staat is het minimum te vinden. De keuze van T_x en T_y moet dus zo goed mogelijk gebaseerd worden op de dominante grootte en richting van de metingen.
- Wanneer direct begonnen wordt met de niet-lineaire solver (hoewel dit wordt afgeraden is dat wel mogelijk), ontstaan er op basis van de rusttoestand direct in de beginsituatie allerlei singulariteiten in de formuleringen van de kostenfunctie en van de gradiënt voor zover deze op richtingen betrekking hebben (immers, aan de nulvektor $\vec{0}$ kan geen éénduidige richting worden toegekend). Hoewel binnen het ISTM in deze situatie is voorzien (kostenfunctie en gradiëntvektor worden identiek aan 0.0 gesteld), lijkt voor de niet-lineaire solver de rusttoestand als startpunt niet optimaal. Door zelf een andere begintoestand te kiezen (naders dan $\vec{T} \equiv \vec{0}$), kan deze situatie worden vermeden.
- Wanneer het niet lukt om met de niet-lineaire solver een (geschikt) minimum te vinden, kan de begintoestand worden aangewend als een 'keuzeknop' om startende vanuit een andere keuze van die begintoestand opnieuw een minimum te gaan berekenen. Het afleiden van een 'dominante' begintoestand vanuit enkel transportmetingen is immers niet éénduidig.

A.2 Voorbeeld van invoerbestand

Onderstaand bestand is een van de invoerbestanden zoals toegepast bij het estuarium-model. Dit model is zoals beschreven in Paragraaf 2.2 en Appendix C van Van den Boogaard (1994).

```

*
* +-----+
* |ISTM-STUURBESTAND dd. 29-10-1999 |
* +-----+
*
* sigma's voor dichte randen, en droge punten:
*
* 0.1e-03      0.1e-03
*
* +-----+
* |Eerste orde regularisaties |
* +-----+
*
* +-----+
* |dTx/dx| dTy/dx | dTx/dy | dTy/dy|
* +-----+
* | V5   | V6   | V7   | V8   |
* +-----+
*
* 0      0      0      0
* 1      1      1      1
* 0.05e-04 0.07e-04 0.05e-04 0.07e-04
*
* +-----+
* |Tweede orde regularisaties |
* +-----+
*
* +-----+
* |d²Tx/dx²| d²Ty/dx²| d²Tx/(dx"dy)| d²Ty/(dx"dy)| d²Tx/dy² | d²Ty/dy²|
* +-----+
* |V9      | V10     | V11      | V12      | V13      | V14      |
* +-----+
*
* 0      0      0      0      0      0
* 1      1      1      1      1      1
* 0.25E-07 0.40E-07 0.17e-07 0.25e-07 0.20e-07 0.30e-07
*
* +-----+
* |Keuzeparameter voor P0 |
* +-----+
*
* = 0 : default (met defaults voor Tx en Ty, anders 1.0)
*       voorlopig dit aanhouden, neg. waarde is disable
* = 1 : P0 inlezen ISTM-bestand
* = 2 : P0 volgens covariantie-polygonen
*
* 0      1.0e-03 1.0e-03
*
* +-----+
* |Invoerbestanden |
* +-----+
*
* +-----+
* | M E E T B E S T A N D E N|
* +-----+
*
* |Bergingsbestand | TRANSPORT-bestand |MODEL-bestand |POLYGOON-
bestand|
* +-----+
*

```

```

*
*      ber_meet.inp      tra_meet.inp      model.dat
** +-----+
* |Uitvoerbestanden |
* +-----+
*
* +-----+
* |               | B E R E K E N I N G |
* +-----+
* |Printbestand   |Bergingsbestand   | TRANSPORT-bestand |
* +-----+
*
*      print.008      ber_tek.008      tra_tek.008
*
* +-----+
* |STAT-bestand   | KOSTEN-bestand   | DIAG-bestand |
* +-----+
*
*      stat.008      cost.008      diag.008
*
*
*
* +-----+
* |Nummering van gridcellen |
* +-----+
*
*   = 1 : lexicografische nummering
*   = 2 : schaakbordnummering
*   = 3 : zebra-nummering
*
*   1
*
* +-----+
* |Lineaire Solver   [SLAP] |
* +-----+
*
*   1  15000  1.E-5
*
* +-----+
* |Niet Lineaire Solver   [1=NEWTON,2=CGM] |
* +-----+
*
*   2  15000  1.E-5
*
* +-----+
* |Niet-lineaire richtingsspreiding |
* +-----+
*
*   1.0
*
* +-----+
* |Startwaarden voor Tx en Ty |
* +-----+
*
*   0      0

```

B Testen van het richtingenformalisme in het niet-lineaire ISTM

B.1 Achtergrond

In dit appendix wordt aan de hand van test berekeningen het effect getoond van het gebruik van het niet-lineaire richtingenformalisme zoals dat in het ISTM is geïmplementeerd volgens de beschrijving van Optie 1 in Hoofdstuk 6. In het bijzonder wordt met deze toepassingen gedemonstreerd dat het niet-lineaire ISTM in staat is om gemeten richtingen $\hat{\phi}$ in strikte zin te reproduceren waar het lineaire model geen onderscheid kan maken tussen die richting en de daarmee overeenkomstige anti-parallelle richting $\hat{\phi} + 180^\circ$. Hierbij wordt opgemerkt dat de doelstelling van de hier beschreven case niet meer dan dat is: afgezien van deze gevoeligheid voor richtingen in strikte zin zijn we in deze toepassingen niet geïnteresseerd in de precieze vorm van de met het ISTM geïdentificeerde transportvelden. Als zodanig wordt er dus geen uitgebreid gevoeligheidsonderzoek uitgevoerd voor variaties in de instellingen van de diverse spreidingen in de meet- en regularisatievergelijkingen om tot een oordeel te komen of het ene transportenveld beter of realistischer is dan een ander veld. Vanuit deze optiek volstaat het dan om uit te gaan van een geschikte artificiële case van relatief beperkte grootte waarin de in te voeren metingen analytisch gekonstrueerd zijn. ‘Geschikt’ betekent hier dat de invoer-data sets zodanig gekozen zijn dat a priori onderscheid mag worden verwacht tussen de performance van het lineaire en niet-lineaire model. Bij een korrekte implementatie van Optie 1 moet dit onderscheid in de diverse ISTM-resultaten worden herkend.

B.2 Beschrijving van de test-case

De test-case is geschematiseerd als een rechthoekig gebied Ω met daarop een rechthoekig $M \times N$ rooster voor de spatiële diskretisatie. Zowel in de x - als de y -richting is het rooster equidistant met vakgroottes Δx en Δy respectievelijk.

T.a.v. van transporten bestaat de metingen-set uit een uniform transport in de x -richting van de vorm $\vec{T}_0 = (T_0, 0)^T$ gegeven voor alle punten van de ‘vertikale’ raaien aan de ‘linker’ en ‘rechter’ rand van het model, namelijk te $m=2$ en $m=M$. De reden dat $m=2$ is gekozen in plaats van $m=1$ heeft te maken met de randafhandeling in staggered grids. De modelranden lopen namelijk over de diepte- ofwel bergingsposities waardoor voor cellen aan de linkerrand (d.w.z. $m=1$) de y -componenten van de transporten formeel buiten het feitelijke rekengebied vallen. Op dezelfde manier vallen de x -componenten van de transporten op de onderrand (d.w.z. $n=1$) eveneens buiten het feitelijke rekengebied. Samenvattend is de meetkundige verzameling van de (m,n) -locaties waar $\vec{T}_0 = (T_0, 0)^T$ als metingen zijn opgegeven dan bepaald door: $V_T = \{(m,n) \mid (m=2 \vee m=M) \wedge 1 \leq n \leq N\}$

Aan de x -component T_0 wordt voor de onzekerheid de spreiding σ_{T_x} toegevoegd terwijl dat voor de y -component wordt gedaan met een spreiding σ_{T_y} .

T.a.v. van transportenrichtingen worden metingen opgegeven voor de centrale verticale raai te $m=(2+M)/2$. Voor de grootte van die richtingen $\hat{\phi}$ langs die raai wordt het volgende voorschrift gehanteerd: voor $T^n := T_0 \cdot \sin(3\pi \cdot \frac{n-1}{N-1})$ groter 0 dan is $\hat{\phi} = 0^\circ$, terwijl voor T^n kleiner dan 0 voor de gemeten richting $\hat{\phi} = 180^\circ$ wordt opgegeven. Is $T^n = 0$ dan wordt er voor die n geen richting als meting opgegeven. Op deze manier is de richting $\hat{\phi}_{\frac{n}{2+M}}^n$ parallel met de transportmetingen voor de intervallen $1 < n < \frac{1}{3}N$ en $\frac{2}{3}N < n < N$ nabij de onder- en bovenrand van Ω , en is die richting anti-parallel voor het daar tussen gelegen interval $\frac{1}{3}N < n < \frac{2}{3}N$. Aan elke zo gedefinieerde $\hat{\phi}$ werd eenzelfde spreiding σ_ϕ toegevoegd voor het weergeven van de onzekerheid in die richting. Op deze manier is het voorschrift $T^n := T_0 \cdot \sin(3\pi \cdot \frac{n-1}{N-1})$ alleen gebruikt om richtingen te definiëren en worden daaraan geen transportgroottes ontleend.

T.a.v. de invoerdata voor bergingen worden twee varianten beschouwd. In de eerste variant worden bergingen voor elke (m,n) -cel opgegeven en alle ter grootte van 0 en met σ_B als de spreiding voor de onzekerheid van deze metingen. In de tweede variant waaraan berekeningen zullen worden uitgevoerd worden bij de invoer geen bergingsgegevens opgegeven.

Voor de numerieke waarden van de hierboven gegeven parameters zijn de volgende instellingen gekozen (in SI-eenheden): $M=10$, $N=25$, $\Delta x = 250$, $\Delta y = 250$, $T_0 = 100$, $\sigma_{T_x} = 1$, $\sigma_{T_y} = 0.1$, $\sigma_B = 0.001$. In de berekeningen werden alleen de eerste orde regularisaties geactiveerd en alle vier met dezelfde spreiding σ_{R_i} . Het probleem is dan een geschikte waarde te vinden voor die σ_{R_i} . Op basis van oriënterende exercities met het niet-lineaire ISTM, waarbij de strategie werd gevolgd dat voor $\sigma_\phi \approx 1^\circ$ een duidelijk effect van de niet-lineaire representatie van richtingen moest kunnen worden herkend t.o.v. de resultaten bij spreidingen σ_ϕ die ordes groter zijn, werd $\sigma_{R_i} = 0.003$ gekozen.

In de hieronder beschreven berekeningen worden dan de resultaten vergeleken die bij diverse variaties van σ_ϕ worden gevonden.

Bij de hierboven geschetste opzet van de test-case worden de volgende opmerkingen geplaatst:

In Figuur B.2.1 is de verzameling van gemeten transporten en transportrichtingen getekend. Deze illustreert het parallel zijn van de transportrichtingen aan de onder en bovenrand van de centrale raai te $m=6$ met de gemeten transporten aan de linker en rechterranden van het model te $m=2$ en $m=10$. In het midden van de centrale raai zijn die richtingen daarentegen anti-parallel. Hoewel de lengte van de vektoren op de centrale raai te $m=6$ net zo lang zijn getekend als die voor de linker en rechterrandsrand moet steeds worden bedacht dat het ISTM geen informatie krijgt over de groottes van transporten op die centrale raai.

Figuur B.2.1 toont het achterliggende fysische idee van de hier gedefinieerde test-case. Dit idee bestaat uit een zeegat met een geul die een achterliggende zeekom met de zee verbindt. In geval van een getijstroming zal het transport in het diepe deel van de geul zeewaarts gericht zijn (midden in de figuur) en over het ondiepe intergetijdedeel landwaarts (links en rechts in de figuur).

Zoals onder Paragraaf B.1 al is opgemerkt zal bij de analyse van de diverse berekeningen alleen gelet worden op het effect van het niet-lineaire mechanisme in het ISTM voor het verwerken van richtingen. De hier genoemde fysische basis van de test case zal dus niet verder worden aangewend om de diverse ISTM resultaten op praktisch realisme te beoordelen.

Voordat er überhaupt een ISTM-berekening heeft plaatsgevonden kan de geschiktheid van deze case voor het toetsen van het ISTM-richtingenformalisme al goed beargumenteerd worden. Immers, met het lineaire ISTM mag als uitkomst voor elke cel het uniforme veld $\vec{T}_0 = (T_0, 0)^T$ worden verwacht (zie Figuur B.2.2 voor een illustratie). Dit veld is namelijk consistent met elke meetvergelijking (transportgroottes, transportrichtingen, en bergingen) en regularisatievergelijking. Vanwege de 180° -dualisme worden de anti-parallelle richtingen in het midden van de raai te $m=6$ niet als tegenstrijdig met de overige transporten-, richtingen en bergingeninformatie gezien. De regularisatie zorgt dan voor een uniforme interpolatie van de (eveneens uniforme) $\vec{T}_0$ -metingen op linker- en rechterranden. Bij het gebruik van Optie 1 voor de richtingen worden in het niet lineaire ISTM de anti-parallelle richtingen echter in strikte zin geïnterpreteerd. Als aan deze metingen voldoende gewicht wordt toegekend (d.w.z. kleine spreiding σ_φ t.o.v. de overige spreidingen) zal het ISTM steeds meer aan deze eis moeten gaan gehoorzamen om velden te vinden die sterk afwijken van het uniforme veld wat voor het lineaire model mag worden verwacht.

B.3 Resultaten van ISTM toepassingen

De verwachting die aan het einde van de vorige paragraaf werd uitgesproken dat met het lineaire ISTM een uniform transportenveld $\vec{T}_0$ wordt gevonden (zie ook Figuur B.2.2) werd door berekeningen bevestigd. Dit resultaat is bovendien niet gevoelig voor de keuzes van de diverse spreidingen en/of het wel of niet gebruiken van gemeten bergingen. Het enige dat telt is dat de regularisatie wordt geactiveerd.

Voor het in beeld brengen van de performance van het niet-lineaire ISTM worden de resultaten van een achttal variaties vergeleken. Dit achttal bestaat uit de combinatie van drie variaties van de spreiding σ_φ voor de richtingen volgens $\sigma_\varphi \in \{10^\circ, 2^\circ, 1^\circ, \frac{1}{10}^\circ\}$, en twee variaties voor de bergingen, namelijk het wel of niet meenemen van deze grootheden hetgeen hier overeenkomt met $\sigma_B \in \{0.001, \infty\}$.

Voor $\sigma_\varphi = 10^\circ$ blijkt dat voor beide σ_B -varianten (binnen de numerieke precisie) het uniforme veld van Figuur B.2.2 wordt gevonden. Blijkbaar is $\sigma_\varphi = 10^\circ$ dermate groot t.o.v. de overige spreidingen dat het niet lineaire ISTM geen (of verwaarloosbaar) gewicht toekent aan de anti-parallelle richtingsinformatie.

De voor $\sigma_\varphi = 2^\circ$ met het niet-lineaire ISTM gevonden transportvelden zijn geïllustreerd in Figuur B.3.1a voor de variant waarbij bergingen worden meegenomen ($\sigma_B = 0.001$), en in Figuur B.3.1b waarbij geen metingen van bergingen zijn opgenomen in de invoergegevens. Op een overeenkomstige manier tonen de Figuren B.3.2a en B.3.2b de resultaten voor $\sigma_\varphi = 1^\circ$, en de figuren B.3.3a en B.3.3b de geïdentificeerde transportenvelden voor $\sigma_\varphi = 0.1^\circ$.

Bij de resultaten van deze subsessievelijke variaties kunnen dan de volgende konklusies en/of opmerkingen worden geplaatst:

Voor beide varianten t.a.v. het wel of niet meenemen van bergingen wordt gevonden dat naarmate de spreiding σ_φ afneemt (ofwel toenemend gewicht dat aan de gemeten richtingen wordt toegekend) het ISTM steeds nauwkeuriger de gemeten richtingen reproduceert. In het bijzonder illustreert dit dat het niet-lineaire ISTM het verschil herkent tussen een richting $\hat{\varphi}$ en de daarmee anti-parallelle richting $\hat{\varphi} + 180^\circ$. Dat was bij deze case zeer zeker niet het geval voor het lineaire ISTM omdat deze voor alle variaties van σ_φ en σ_B het uniforme transportenveld van Figuur B.2.2 produceert.

Bij de overgang van $\sigma_\varphi = 2^\circ$ naar $\sigma_\varphi = 1^\circ$ is het effect op het transportenveld (en met name de gevoeligheid voor de richtingsinformatie $\hat{\varphi}$) veel groter voor de variant waarbij zonder bergingen wordt gerekend dan voor de variant waarbij de bergingen wel in de invoer zijn opgenomen. Dit is goed te verklaren uit het feit dat in het laatste geval deze bergingen (die alle 0 zijn) in feite als extra regularisatie functioneren.

Als vervolg van het bovenstaande aandachtspunt kan worden gezegd dat het beeld dat de bergingen regulariserend werken ook kan worden herkend in de overgang van $\sigma_\varphi = 1^\circ$ naar $\sigma_\varphi = 0.1^\circ$. Immers, het transportenveld $\vec{T}(\cdot, \cdot)$ van Figuur B.3.3a (inclusief bergingen) heeft over het gehele gebied gezien een geleidelijker spatiaal verloop dan het veld $\vec{T}(\cdot, \cdot)$ van Figuur B.3.3b (exclusief bergingen). In dat laatste veld zijn namelijk zeer sterke veranderingen in de $\vec{T}(\vec{r})$ te zien voor (en vrijwel beperkt tot) de centrale locaties $\vec{r}$ (d.w.z. in de omgeving van het punt $m=6$ en $n=13$) waar de gemeten richtingen tegengesteld zijn aan alle overige gemeten richtingen en transporten.

Zoals hierboven al werd gekonstateerd worden in het traject van variaties $\sigma_\varphi = 2^\circ \rightarrow \sigma_\varphi = 1^\circ \rightarrow \sigma_\varphi = 0.1^\circ$ zowel voor de $\sigma_B = 0.001$ als $\sigma_B = \infty$ variant de gemeten richtingen steeds beter door het ISTM gereproduceerd.

Voor het reproduceren van de gemeten transportvectoren (d.w.z. *groottes*) op de raaien $m=2$ en $m=10$ aan de randen van het modelgebied is dat voor de $\sigma_B = 0.001$ versie duidelijk niet het geval (zie de Figuren B.3.1a, B.3.2a, B.3.3a). De ISTM-schattingen daarvan worden steeds slechter, met name rond het centrum $n=13$ van die twee raaien. Dit is het gevolg van de 1^e orde regularisaties. Deze streven naar een zo uniform mogelijk transportenveld $\vec{T}(\cdot, \cdot)$. Echter, voor afnemende σ_φ moet in toenemende mate eveneens voldaan worden aan de $\hat{\varphi} = 180^\circ$ metingen hetgeen niet consistent is met een uniforme $\vec{T}(\cdot, \cdot)$. De grootste niet-uniformiteit van het uiteindelijk gevonden veld ontstaat dan met name in het centrum van het gebied rond $(m,n)=(6,13)$ waar die $\hat{\varphi} = 180^\circ$ metingen zich bevinden. De regularisatie blijkt echter sterk genoeg om binnen de horizontale strook $6 < n < 19$ zelfs tot aan de randen te $m=2$ en $m=10$ een groot effect uit te oefenen met zware onderschattingen van de daar aanwezige transportgroottes als resultaat.

Dit effect van het slechter reproduceren van de gemeten transportgroottes aan de randen wordt in veel mindere mate gezien voor de $\sigma_B = \infty$ serie (zie de Figuren B.3.1b, B.3.2b, B.3.3b). Enerzijds is dit het gevolg van een minder zware regularisatie omdat het deel van

de regularisatie die bij $\sigma_B = 0.001$ wordt geïnduceerd door de gemeten bergingen nu afwezig is. Minstens zo belangrijk is echter dat bij de $\sigma_B = 0.001$ serie de door de gemeten bergingen geïnduceerde regularisatie (in de vorm van de continuïteitsvergelijking met nul als rechterlid) *niet consistent* is met de verzameling gemeten transportgroottes en transportrichtingen. Met het weglaten van de bergingen is dan het door deze inkonsistentie veroorzaakte effect op het te schatten veld $\vec{T}(\cdot, \cdot)$ eveneens afwezig. Hierbij is het wel opvallend dat bij de berekeningen zonder bergingen de transportgroottes op de randen $m=2$ en $m=10$ zeer goed gereproduceerd worden voor $\sigma = 2.0^\circ$ en $\sigma = 0.1^\circ$ terwijl dat minder is voor de tussenliggende $\sigma = 1.0^\circ$. Hiervoor is op dit moment nog geen afdoende verklaring voor handen.

Een zeer opvallend aspect in de diverse figuren is dat de met het ISTM gevonden transportvelden $\vec{T}(\cdot, \cdot)$ niet symmetrisch zijn t.o.v. de centrale horizontale raai te $n=13$. Deze symmetrie werd van te voren wel verwacht vanwege de symmetrie van de invoergegevens. Uit aanvullende berekeningen waarbij de optimalisatie werd herhaald voor een alternatieve initial guess van het transportenveld bleek de verwachte symmetrie zich echter op een andere manier te manifesteren. Alvorens de resultaten van die additionele berekeningen te laten zien worden kort de resultaten samengevat van een meer theoretische evaluatie van de symmetrie in de huidige case. In deze theoretische evaluatie werd gekeken naar de symmetrie van de diverse componenten van de totale kostenfunctie die in in deze case geminimaliseerd moet worden. Deze componenten bestaan uit de bijdragen van (i) gemeten transportgroottes, (ii) gemeten transportrichtingen, (iii) vier 1^e-orde regularisatietermen en (iv) de bergingen. Zonder op de details en/of afleidingen in te gaan wordt dan het volgende gevonden, waarbij we onderscheid maken t.a.v. het wel of niet meennemen van bergingen (d.w.z. component iv).

Als er *geen bergingen worden meegenomen in de invoer* kan worden aangetoond dat als $\vec{T} = (T_x, T_y)^T$ een oplossing is (d.w.z. een minimum van de totale kostenfunctie) dat dan het veld $\vec{T}^*$ gedefinieerd volgens,

$$\begin{cases} (T_x^*)^n = (T_x)^n_m \\ (T_y^*)^n = -(T_y)^n_m \end{cases} \quad (\text{B.3.1})$$

dat ook is. Vergelijking (B.3.1) laat zien dat het $\vec{T}^*$ veld uit $\vec{T}$ wordt gevonden door voor elke positie het teken van de T_y component te inverteren. Dat naast $\vec{T}$ ook $\vec{T}^*$ een oplossing is volgt uit de symmetrieën van zowel de meetgegevens als de gebruikte regularisaties. In algemene praktijktoepassingen zullen al deze eigenschappen meestal niet gelden.

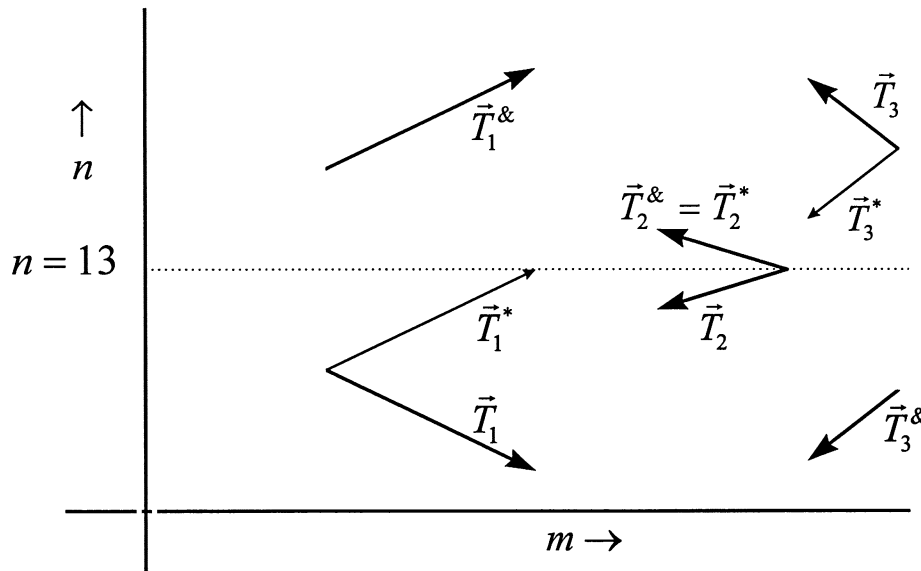
Als er *wel bergingen worden meegenomen in de invoer* dan geldt bovenstaande eigenschap dat met $\vec{T}$ ook $\vec{T}^*$ een oplossing is niet meer. De symmetrie-eigenschap wordt immers verbroken door de kostenfunctie J_B van de bergingen die met $\hat{B}_m^n = 0$ hier de volgende vorm heeft:

$$J_B(\vec{T}) = \frac{1}{\sigma_B^2} \sum_{m,n} \left(\frac{(T_x)^n_m - (T_x)^{n-1}_{m-1}}{\Delta x} + \frac{(T_y)^n_m - (T_y)^{n-1}_m}{\Delta y} \right)^2 \quad (\text{B.3.2})$$

In het algemeen (d.w.z. niet alle $(T_y)_m^n$ zijn 0) zullen dan $J_B(\vec{T})$ en $J_B(\vec{T}^*)$ verschillend zijn. In plaats van $\vec{T}^*$ kan er echter nu een andere alternatieve oplossing uit $\vec{T}$ gekonstrueerd worden. Het blijkt namelijk dat als $\vec{T} = (T_x, T_y)^T$ een oplossing is (d.w.z. een minimum van de totale kostenfunctie, inclusief de J_B van vergelijking B.3.2) dat dan dat ook geldt voor het veld $\vec{T}^\&$ gedefinieerd volgens:

$$\begin{cases} (T_x^\&)_m^n = (T_x)_m^{26-n} = (T_x^*)_m^{26-n} \\ (T_y^\&)_m^n = -(T_y)_m^{26-n} = -(T_y^*)_m^{26-n} \end{cases} \quad (\text{B.3.3})$$

Het $\vec{T}^\&$ veld wordt in twee stappen uit $\vec{T}$ gevonden. De eerste stap bestaat uit het voor elke (m,n) -positie inverteren van het teken van de T_y component van $\vec{T}$. Dit levert dan het $\vec{T}^*$ veld van vergelijking B.3.1 op. In de tweede stap worden de (m,n) -posities waar de $\vec{T}^*$ -vectoren aangrijpen gespiegeld in de horizontale raai te $n=13$. De $\vec{T}^*$ vectoren worden bij die spiegeling in ongewijzigde vorm meegesleept. Beide stappen samen nemend betekent dit dat de vektor in zijn geheel, dus van aangrijpingspunt tot en met eindpunt van de pijl, gespiegeld wordt in de lijn $n=13$. In een figuur ziet de transformatie $\vec{T} \rightarrow \vec{T}^* \rightarrow \vec{T}^\&$ er dan als volgt uit:



Hoewel het $\vec{T}^\&$ veld hier is geïntroduceerd voor het geval dat er inclusief bergingen wordt gerekend moet worden opgemerkt dat ook voor het speciale geval dat er geen bergingen worden meegenomen dit veld naast $\vec{T}$ een additionele oplossing is van het optimalisatieprobleem. Deze $\vec{T}^\&$ kan dan echter samenvallen met het eerder beschouwde $\vec{T}^*$ veld. Ook nu moet worden vermeld dat deze bevindingen strikt gerelateerd zijn aan de hier gedefinieerde artificiële/synthetische ISTM-case, en meestal niet geldig zullen zijn in meer algemene praktijktoepassingen.

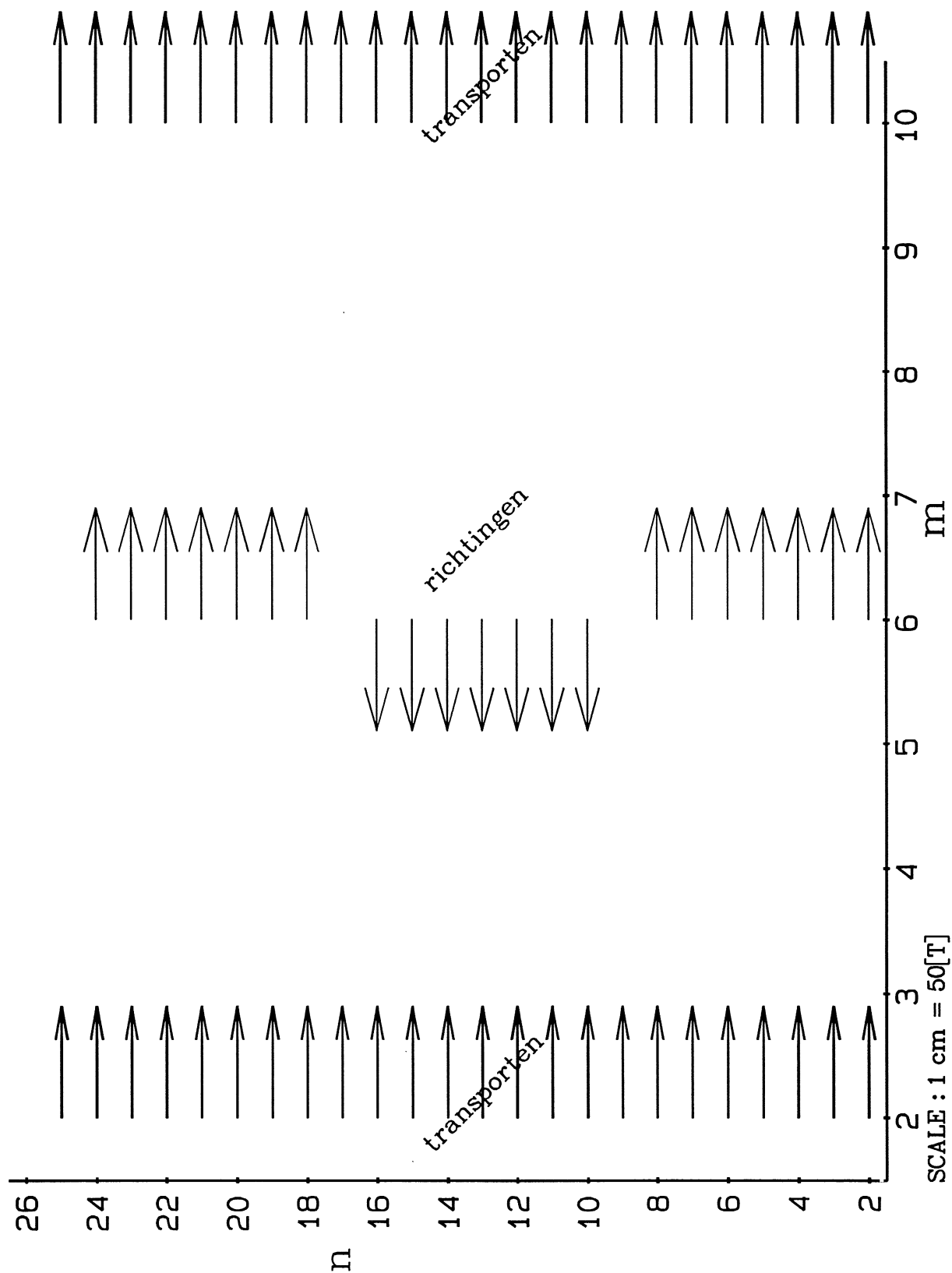
Om bovenstaande bevindingen te illustreren aan de hand van ISTM-berekeningen zijn twee van de hierboven beschreven runs opnieuw uitgevoerd maar nu werd als initial guess het uniforme veld $\vec{T}_{ini} = (T_0, 1)^T$ gekozen in plaats van $\vec{T}_{ini} = (T_0, 0)^T$ zoals voorheen (hierbij moet worden vermeld dat bij een initialisatie volgens $\vec{T}_{ini} = (T_0, -1)^T$, of een andere negatieve waarde voor $(T_y)_{ini}$, de eerder getoonde resultaten werden teruggevonden)

In het eerste experiment werd gerekend met $\sigma_\varphi = 1.0^\circ$ en $\sigma_B = 0.001$, dus *inclusief bergingen* in de invoer data set. Het resultaat van het dan met het ISTM gevonden transportenveld is te zien in Figuur B.3.4 en moet vergeleken worden met dat van Figuur B.3.2a. Het valt dan meteen op dat deze velden niet hetzelfde zijn: de pijlen wijzen in Figuur B.3.2a vrijwel allemaal naar beneden (d.w.z. in de negatieve n -richting) terwijl ze in Figuur B.3.4 naar boven wijzen. Het gevonden resultaat is dus duidelijk afhankelijk van de beginschatting $\vec{T}_{ini}$. Wel zijn deze velden sterk gerelateerd qua symmetrie hetgeen we aan de hand van bovenstaande theoretische beschouwing nu expliciet kunnen toetsen. Dit is als volgt gedaan. Uit het $\vec{T}$ -veld van Figuur B.3.4 zijn conform het voorschrift van de vergelijkingen (B.3.1) en (B.3.3) de velden $\vec{T}^*$ en/of $\vec{T}^\&$ gekonstrueerd en deze zijn vervolgens vergeleken met de $\vec{T}$ van Figuur B.3.2a. In Figuur B.3.5a zijn zowel de $\vec{T}$ van Figuur B.3.2 als het $\vec{T}^*$ veld getekend, terwijl in Figuur B.3.5b $\vec{T}$ en $\vec{T}^\&$ tegelijk worden getoond. Dan blijkt dat $\vec{T}^*$ enigszins verschillend is van de $\vec{T}$ van Figuur B.3.2, terwijl deze $\vec{T}$ zo goed als samenvalt met $\vec{T}^\&$. Dit is precies zoals te verwachten uit bovenstaande theoretische beschouwing t.a.v. het bestaan van meervoudige oplossingen en de relatie die die oplossingen dan moeten hebben. Deze beschouwing leerde namelijk dat voor dit experiment $\vec{T}^*$ niet, en $\vec{T}^\&$ wel een alternatieve oplossing vormen.

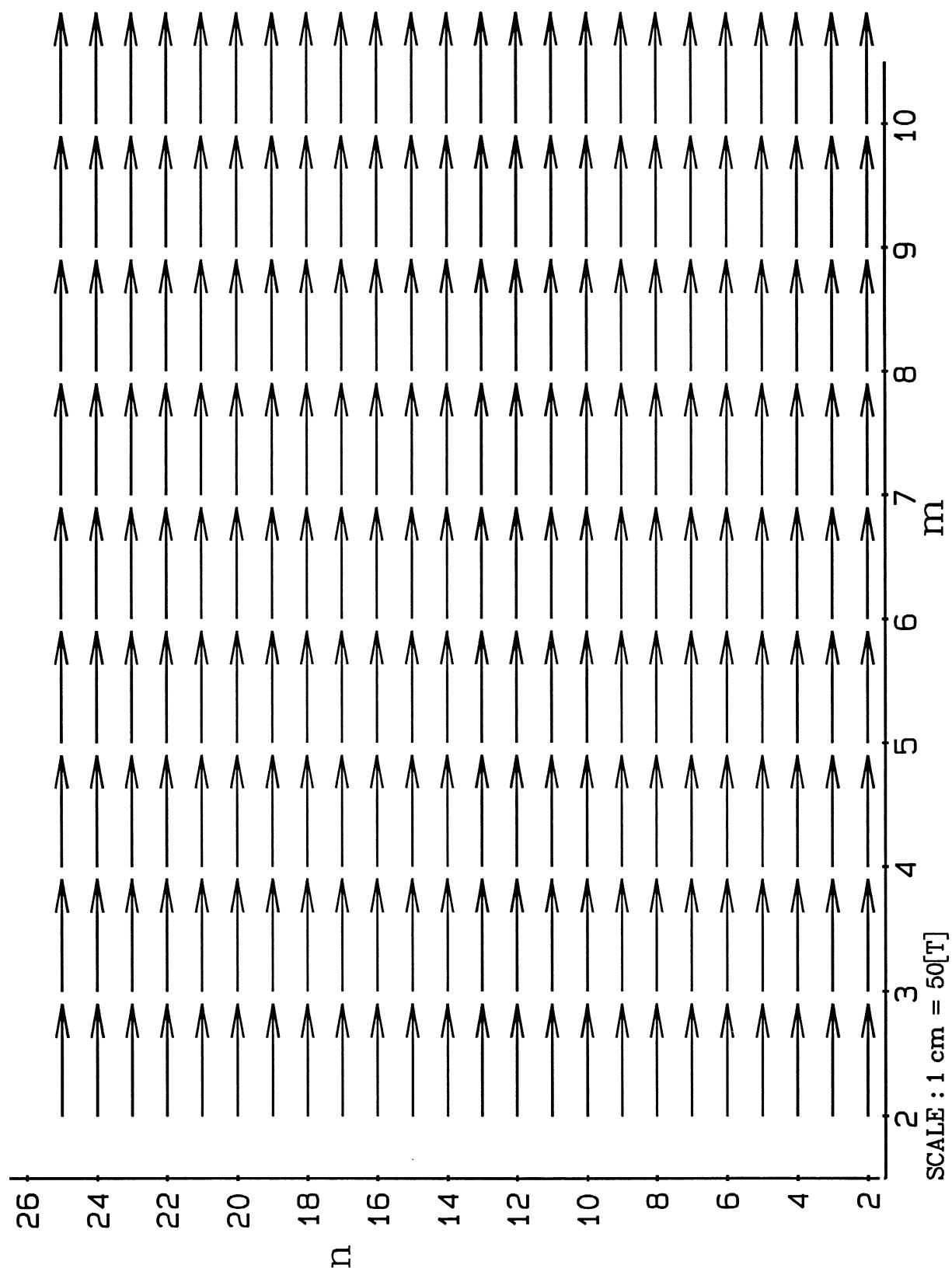
In het tweede experiment werd gerekend met $\sigma_\varphi = 0.1^\circ$ en $\sigma_B = \infty$, dus *exclusief bergingen* in de invoer data set. Het resultaat van het dan met het ISTM gevonden transportenveld is te zien in Figuur B.3.6 en deze $\vec{T}$ moet dan vergeleken worden met het veld van Figuur B.3.3b. Weer valt op dat deze velden verschillend zijn maar wel in sterke mate gerelateerd/symmetrisch. Dit is op dezelfde manier getoetst als in het vorige experiment door uit het $\vec{T}$ -veld van Figuur B.3.6 de velden $\vec{T}^*$ en/of $\vec{T}^\&$ te konstrueren en deze te vergelijken met de $\vec{T}$ van Figuur B.3.3b. Deze $\vec{T}$ en $\vec{T}^*$ zijn te zien in Figuur B.3.7a, terwijl $\vec{T}$ én $\vec{T}^\&$ te vinden zijn in Figuur B.3.7a. Nu is er geen verschil te zien tussen $\vec{T}^*$ en $\vec{T}^\&$ en beide velden komen goed overeen met het veld $\vec{T}$ van Figuur B.3.3b (de iets grotere verschillen die hier worden gevonden t.o.v. het vorige experiment representeren slechts afbreekeffecten in de optimalisatie-procedure van het niet-lineaire ISTM). Deze resultaten zijn weer conform de hierboven weergegeven theoretische achtergrond t.a.v. het bestaan van meervoudige oplossingen.

Samenvattend laat het voorafgaande zien dat het niet-lineaire ISTM veel dwingender rekening houdt met richtingsinformatie dan het lineaire ISTM. Uitgaande van het fysische idee achter de hier uitgevoerde test case (namelijk een zeegat met een geul die een achterliggende zeekom met de zee verbindt, en waarbij in geval van getijstroming het transport in het diepe gedeelte van de geul zeewaarts gericht zal zijn, en over het ondiepe gedeelte landwaarts) en de op die basis gegenereerde invoer data set hebben de met het niet-lineaire ISTM model gegenereerde transportpatronen een veel groter realisme dan het uniforme veld volgens het lineaire ISTM. Wel moest voor het niet-lineaire ISTM een geschikte keuze voor de spreidingen van de diverse ruizen worden gezocht. De hier

getoonde figuren laten namelijk zien dat de resultaten voor het uiteindelijk gevonden veld $\vec{T}$ zeer gevoelig kunnen zijn voor die keuze. Dat geldt niet alleen voor de σ_φ maar tevens voor de spreidingen van de regularisaties. Bovendien laat het hier getoonde voorbeeld zien dat ook de initial guess van het transportenveld in belangrijke mate het eindresultaat kan beïnvloeden. Deze bevindingen demonstreren dat de hier gebruikte data set zowel kwantitatief als kwalitatief niet voldoende omvangrijk is voor een goed gesteld invers probleem. Die gesteldheid zou bijvoorbeeld al aanzienlijk beter zijn indien op de centrale raai naast richtingsinformatie eveneens groottes van transporten beschikbaar zouden zijn. Voor praktijktoepassingen betekent dit alles dat telkens de nodige voorzichtigheid moet worden betracht bij het opstellen van de doelstelling van een case en diens haalbaarheid gegeven de kwaliteit en kwantiteit van de invoergegevens.



Analytische test case voor richtingen, 10x25 model Invoergegevens transporten en transport-richtingen	Non-Linear ISTM	Nov. 1999
	Meetset Transporten en Richtingen	
DELFT HYDRAULICS	Z2727.30	Fig. B.2.1



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-identificeerd uniform transportenveld

Non-Linear
 ISTM

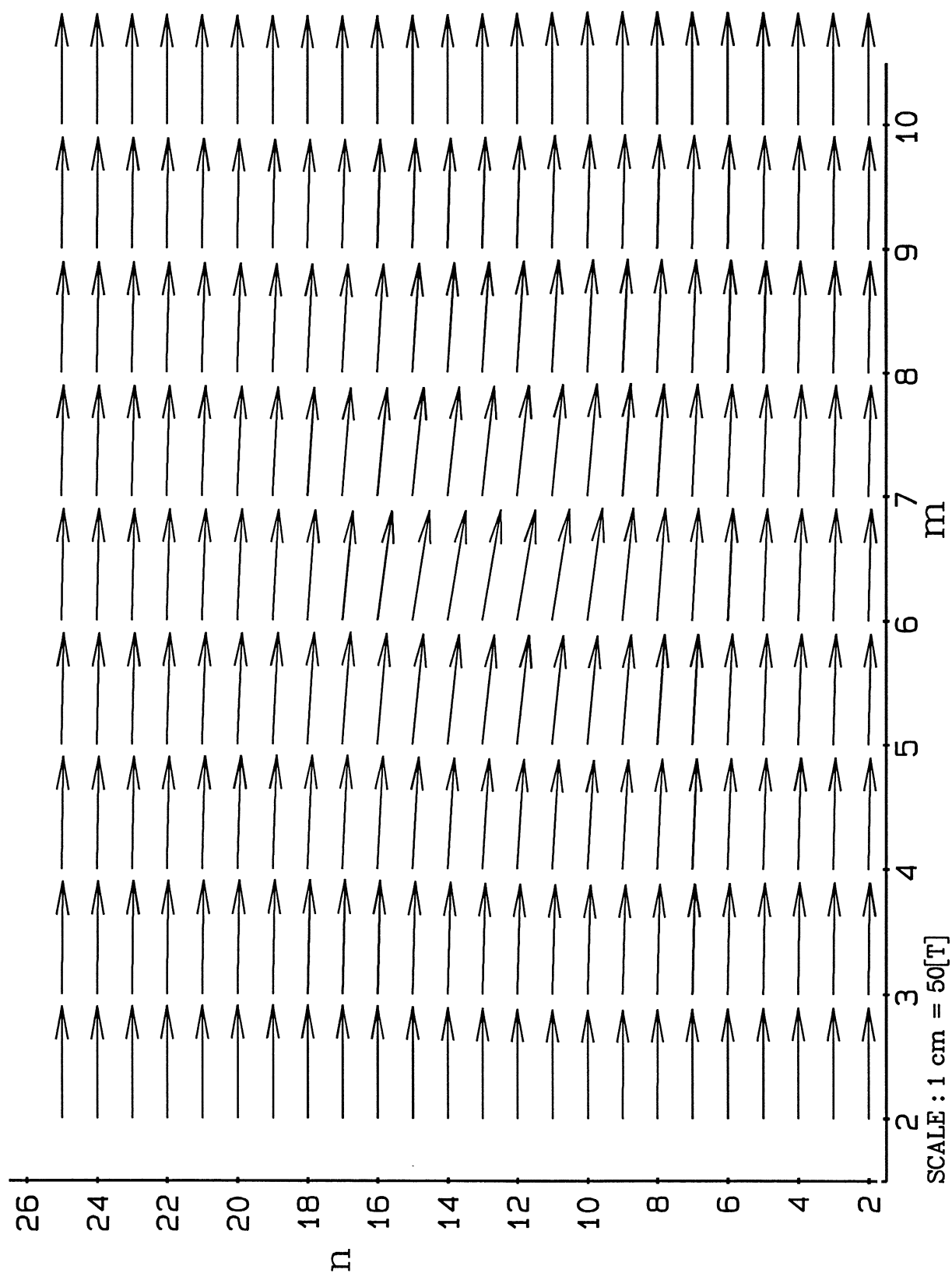
Nov. 1999

Uniform transportenveld

DELFT HYDRAULICS

Z2727.30

Fig. B.2.2



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld

Non-Linear
ISTM

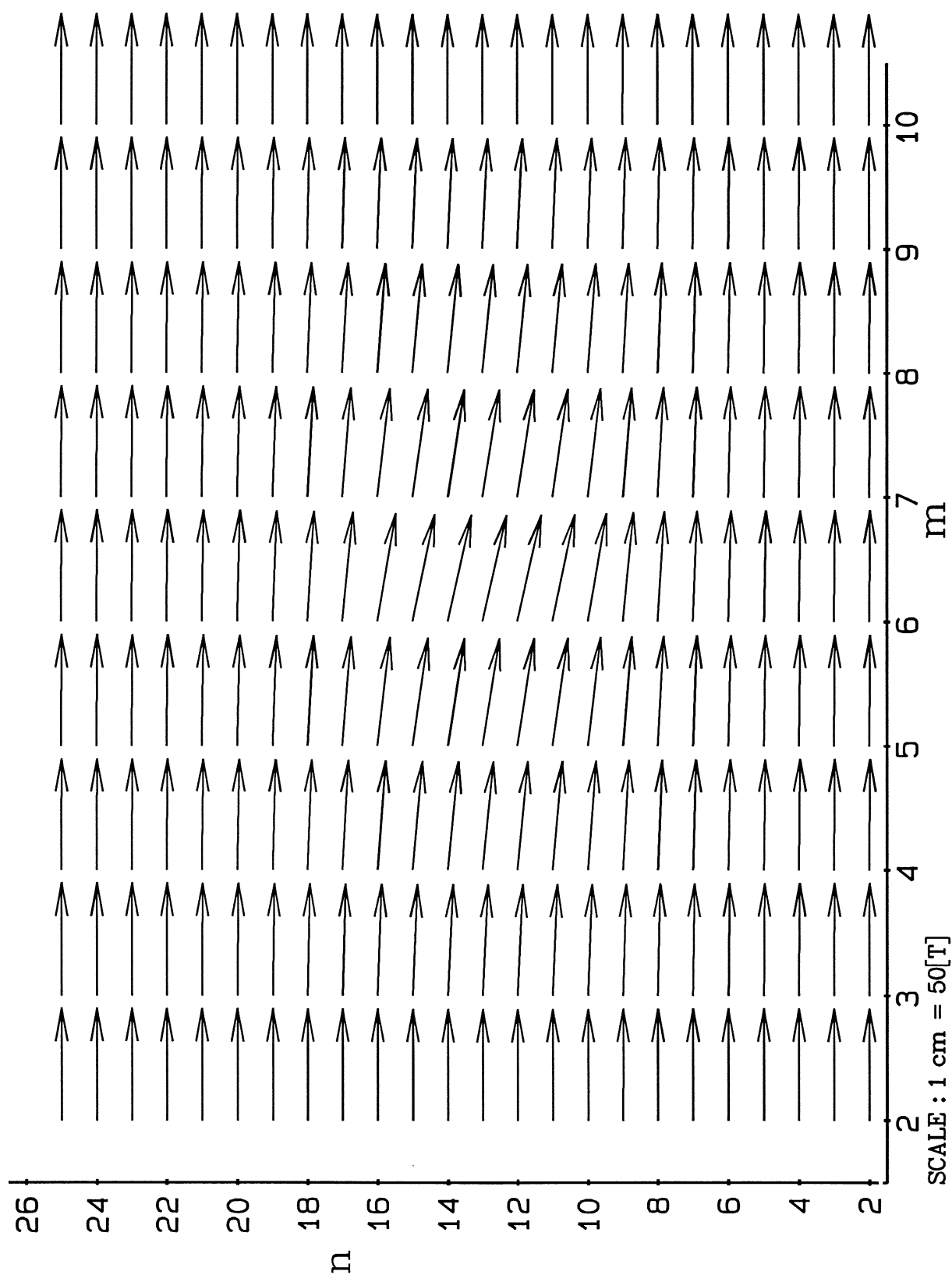
Nov. 1999

Inkl. B., $\sigma_{\varphi} = 2.0^{\circ}$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.1a



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld

Non-Linear
ISTM

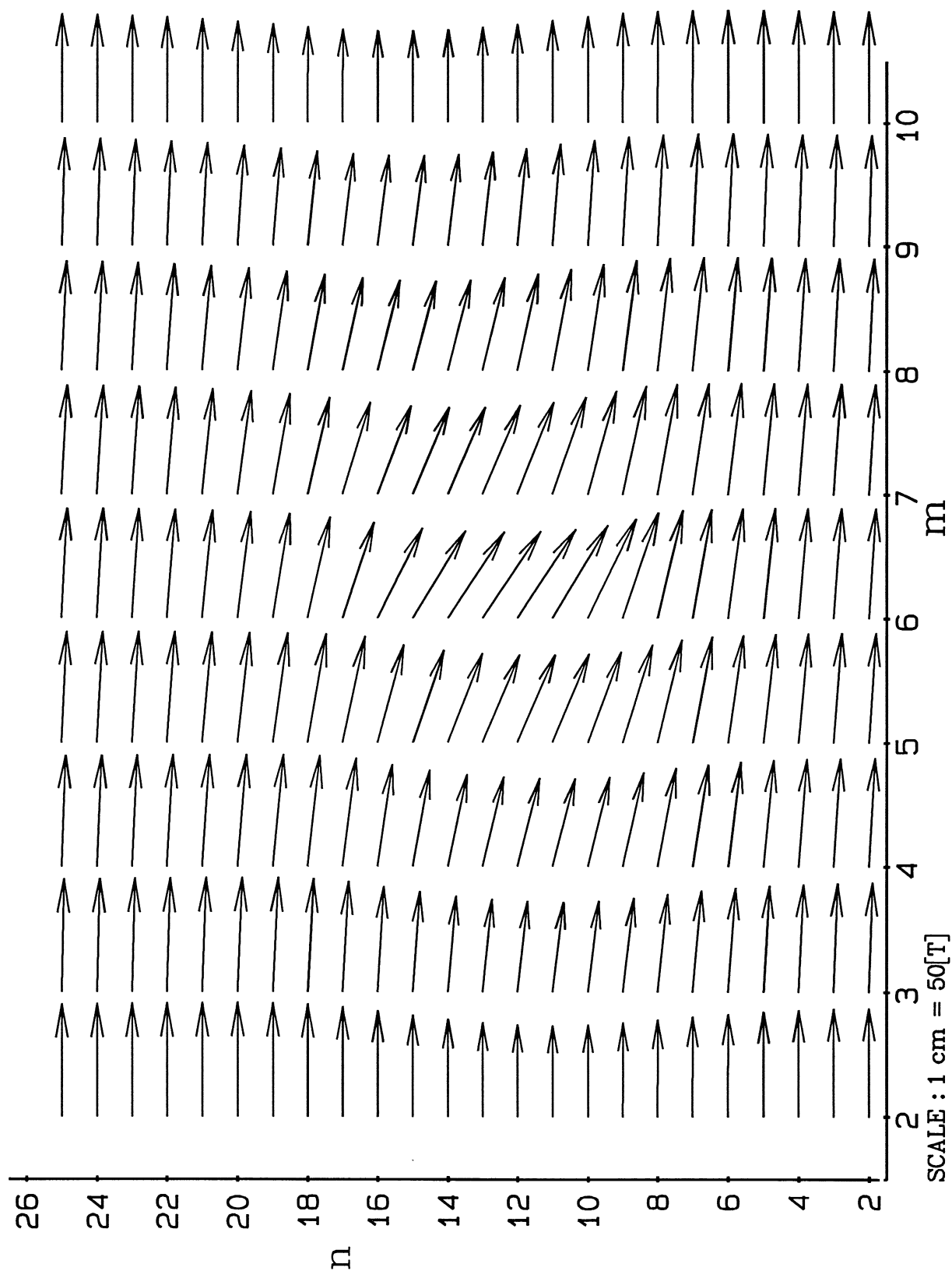
Nov. 1999

Exkl. B., $\sigma_\varphi = 2.0^\circ$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.1b



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld

Non-Linear
ISTM

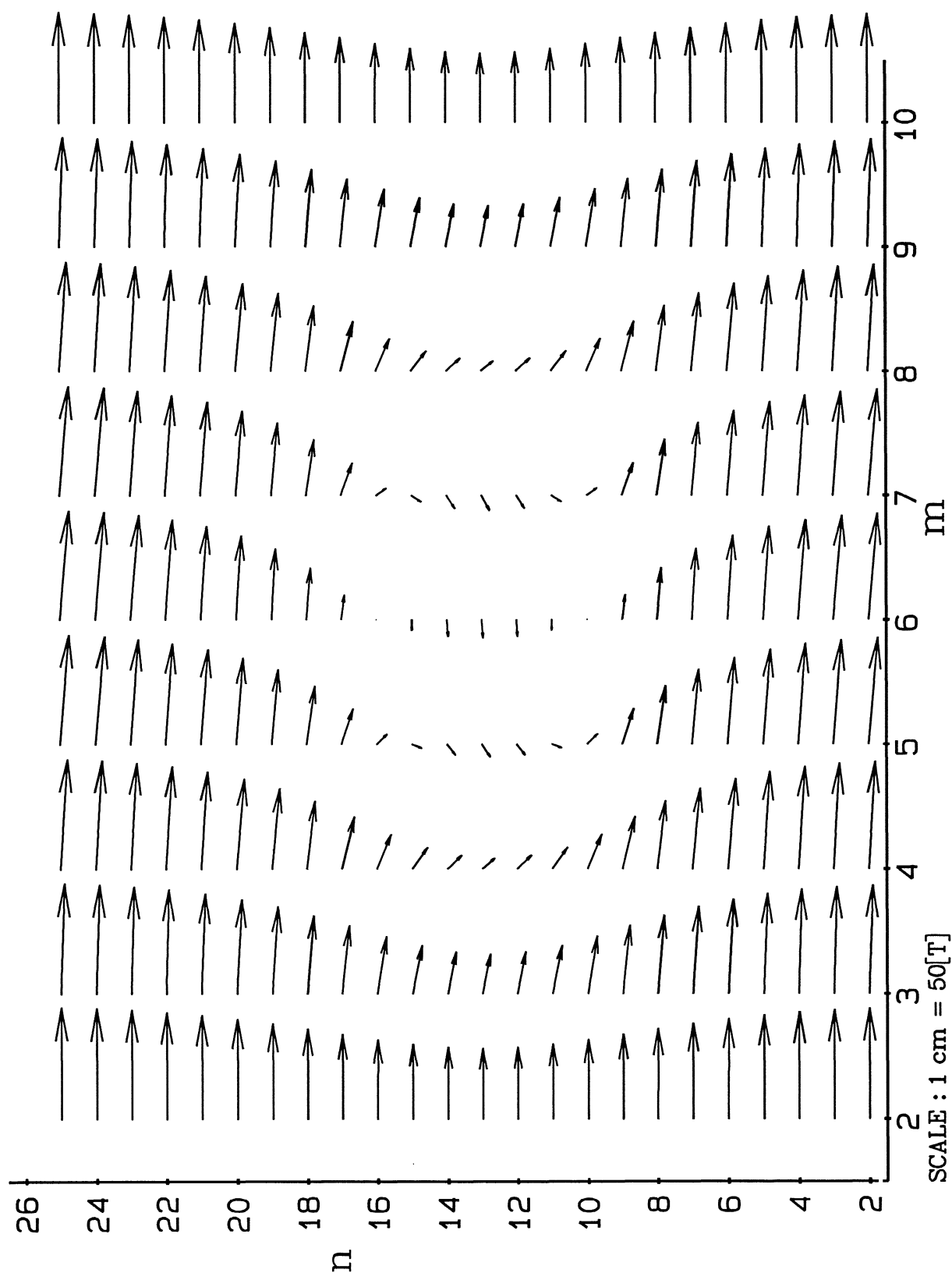
Nov. 1999

Inkl. B., $\sigma_{\varphi} = 1.0^{\circ}$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.2a



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld

Non-Linear
ISTM

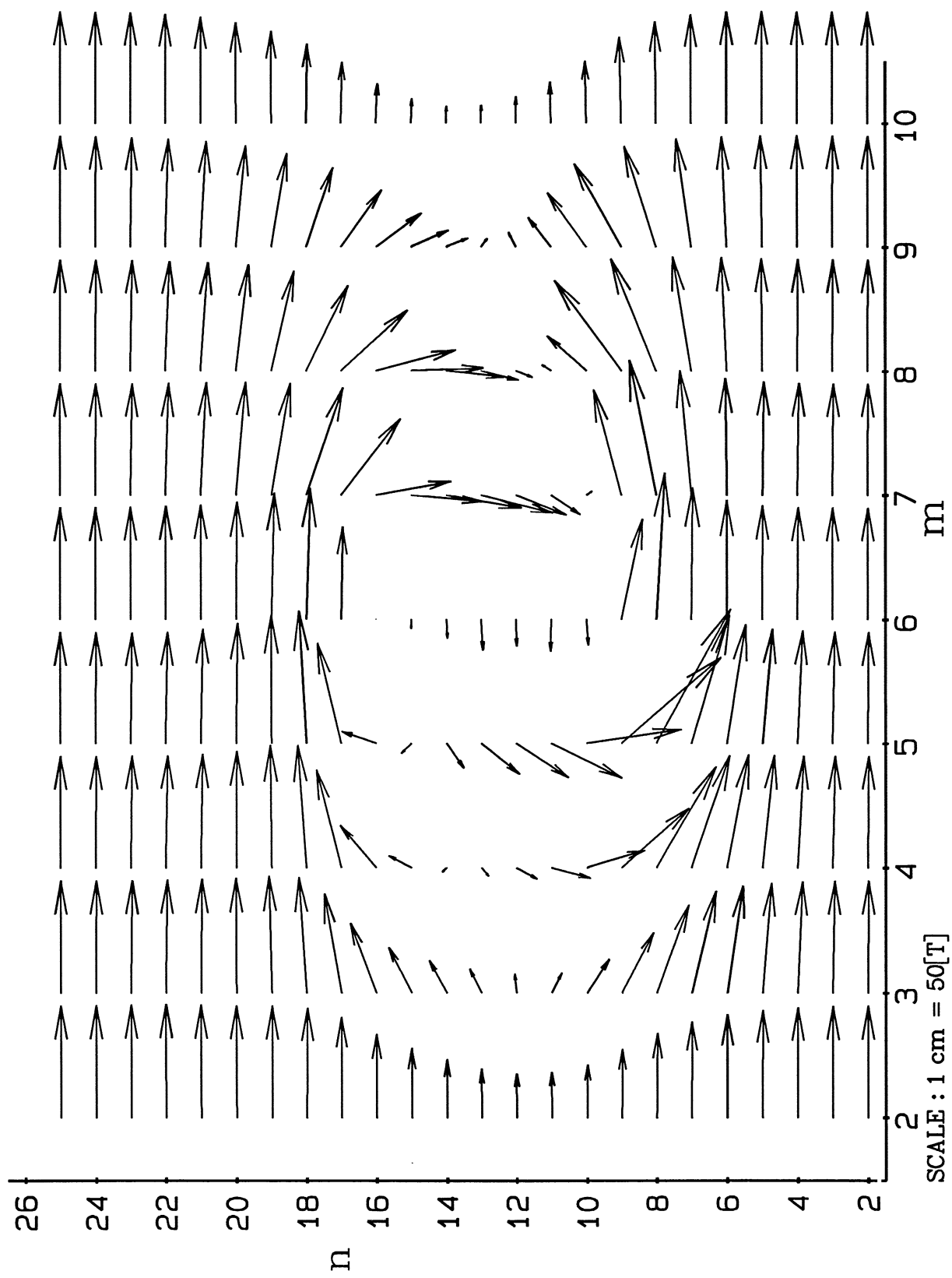
Nov. 1999

Exkl. B., $\sigma_{\varphi} = 1.0^{\circ}$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.2b



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerd transportenveld

Non-Linear
 ISTM

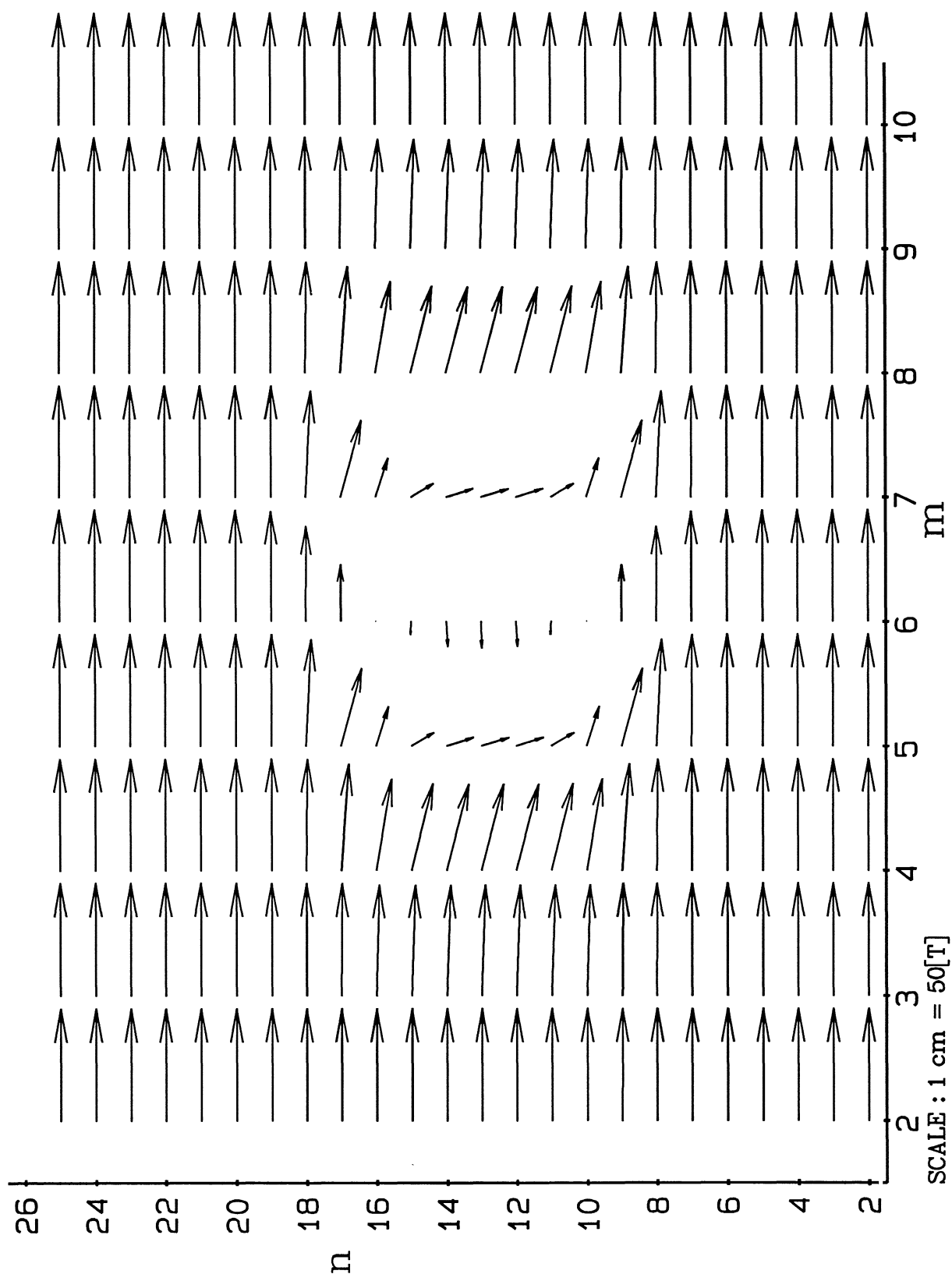
Nov. 1999

Inkl. B., $\sigma_\varphi = 0.1^\circ$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.3a



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld

Non-Linear
ISTM

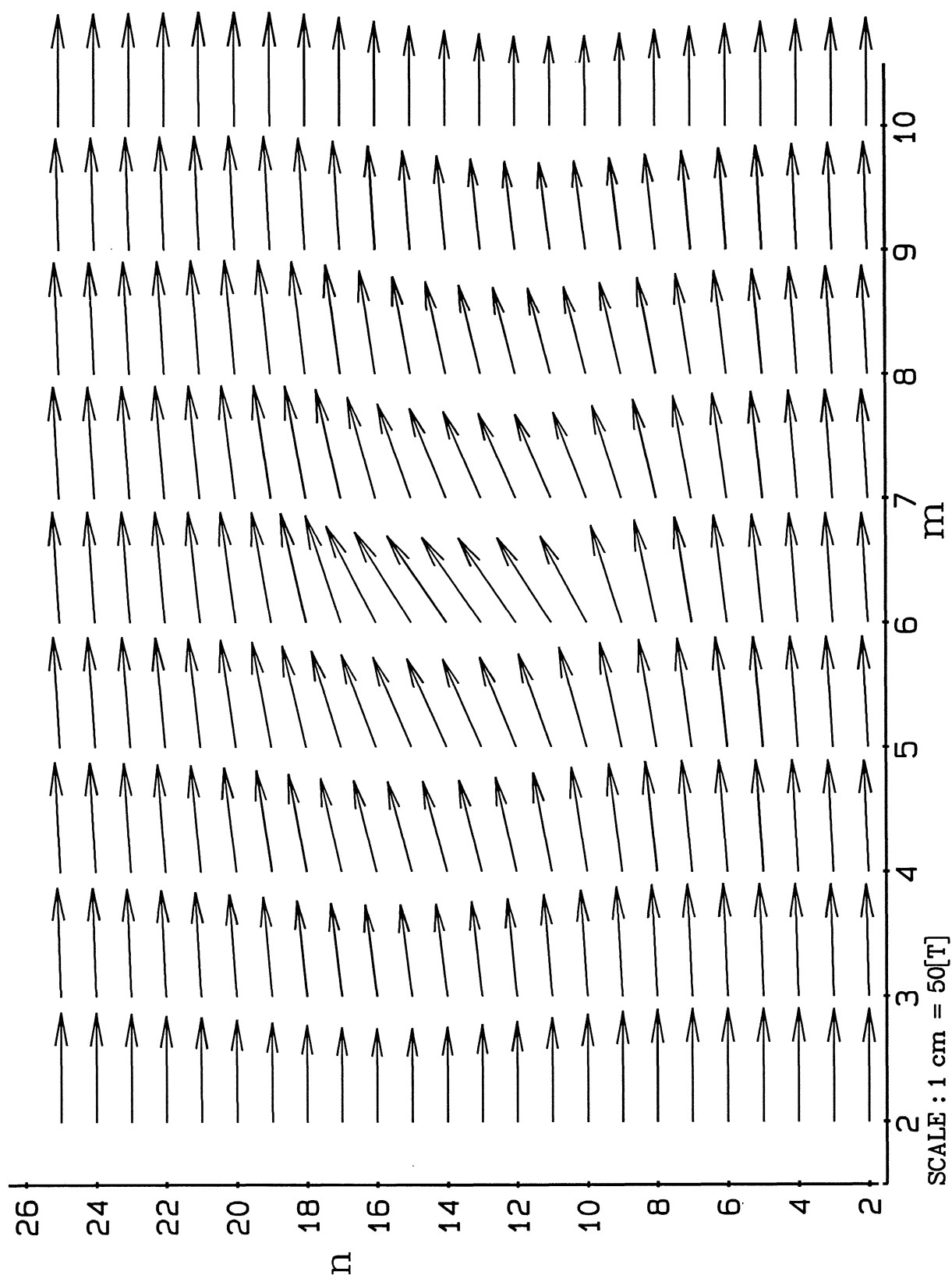
Nov. 1999

Exkl. B., $\sigma_{\phi} = 0.1^{\circ}$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.3b



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerd transportenveld.
 Alternatieve initial guess. Zie ook Figuur B.3.2a

Non-Linear
 ISTM

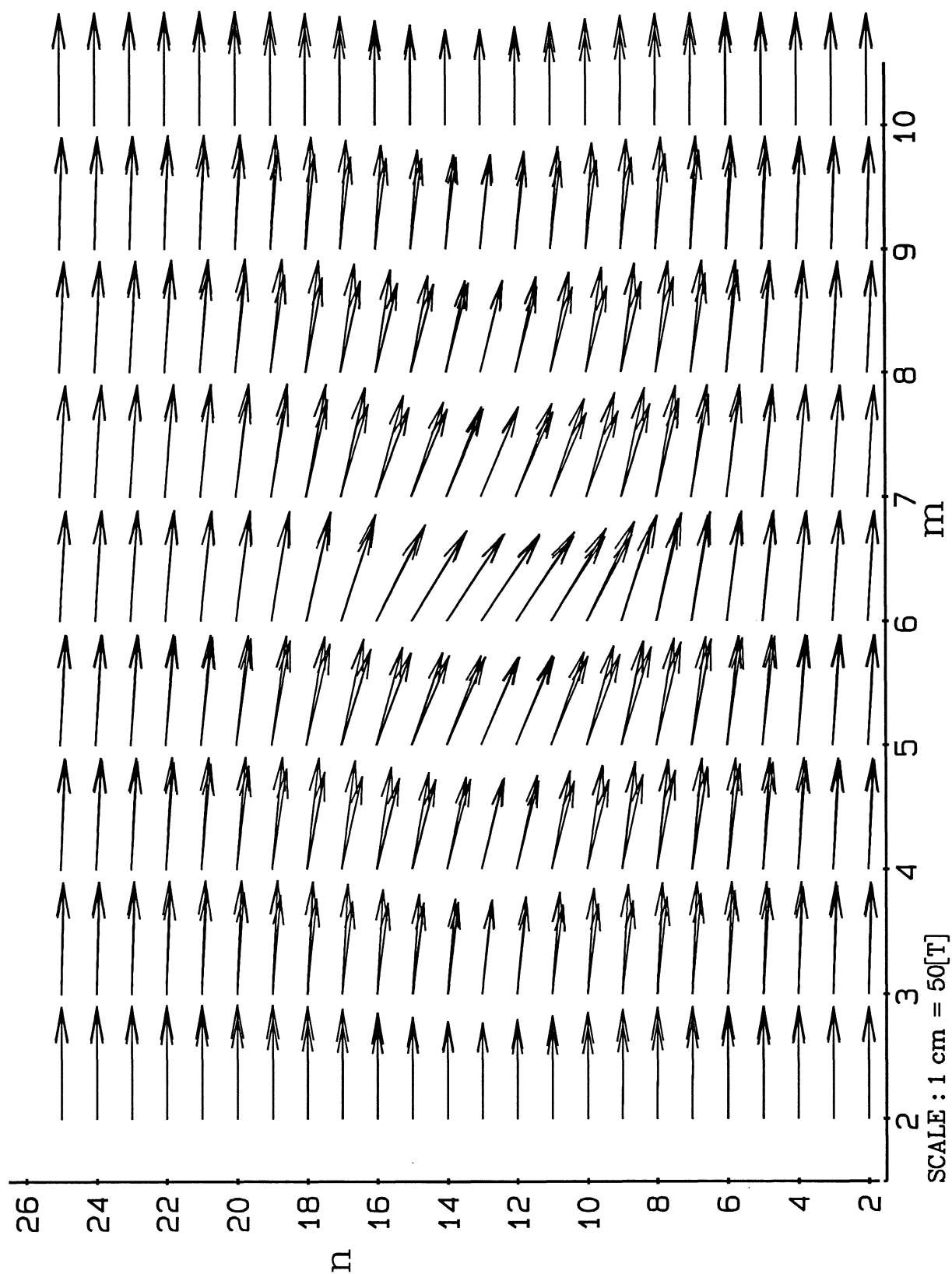
Nov. 1999

Inkl. B., $\sigma_\varphi = 1.0^\circ$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.4



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerde transportenvelden.
 T-veld van Figuur B.3.2a en T*-veld van Figuur B.3.4

Non-Linear
 ISTM

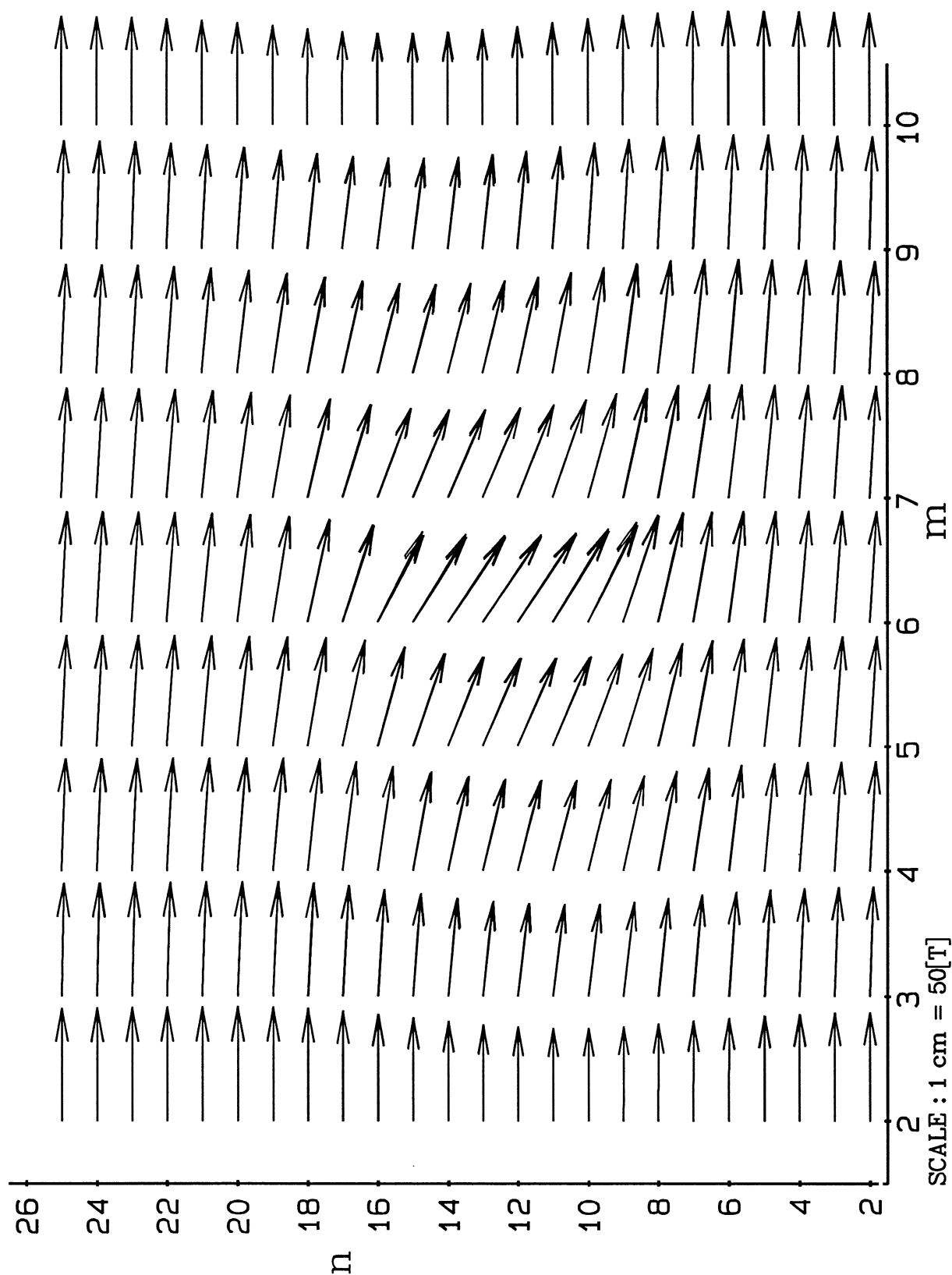
Nov. 1999

Inkl. B., $\sigma_{\varphi} = 1.0^{\circ}$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.5a



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerde transportenvelden.
 T-veld van Figuur B.3.2a en $T^{\&}$ -veld van Figuur B.3.4

DELFT HYDRAULICS

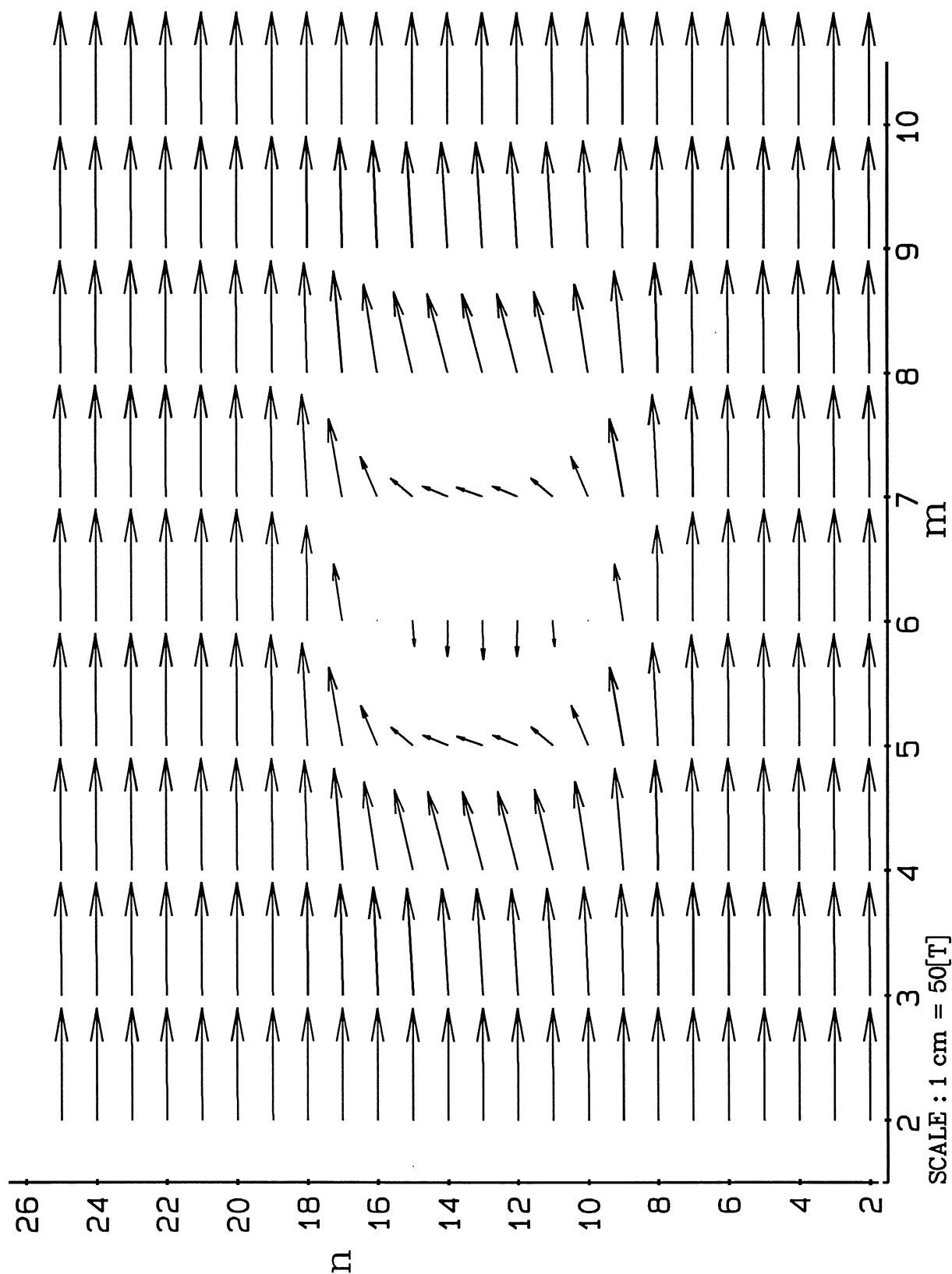
Non-Linear
 ISTM

Inkl. B., $\sigma_{\varphi} = 1.0^{\circ}$

Z2727.30

Nov. 1999

Fig. B.3.5b



Analytische test case voor richtingen, 10x25 model
Met ISTM ge-indentificeerd transportenveld.
Alternatieve initial guess. Zie ook Figuur B.3.3b

DELFT HYDRAULICS

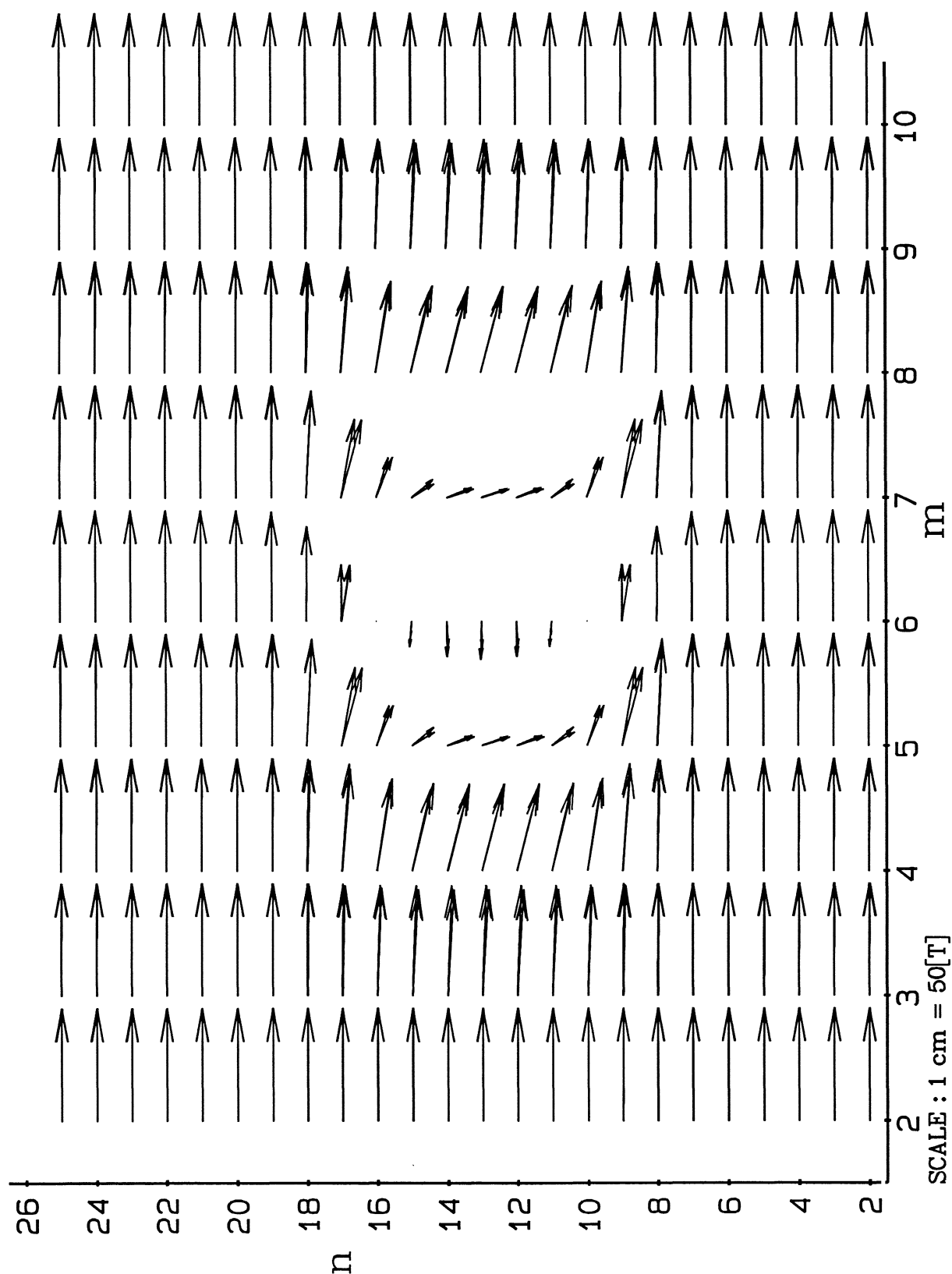
Non-Linear
ISTM

Nov. 1999

Exkl. B., $\sigma_{\varphi} = 0.1^{\circ}$

Z2727.30

Fig. B.3.6



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerde transportenvelden.
 T-veld van Figuur B.3.3b en T^* -veld van Figuur B.3.6

Non-Linear
 ISTM

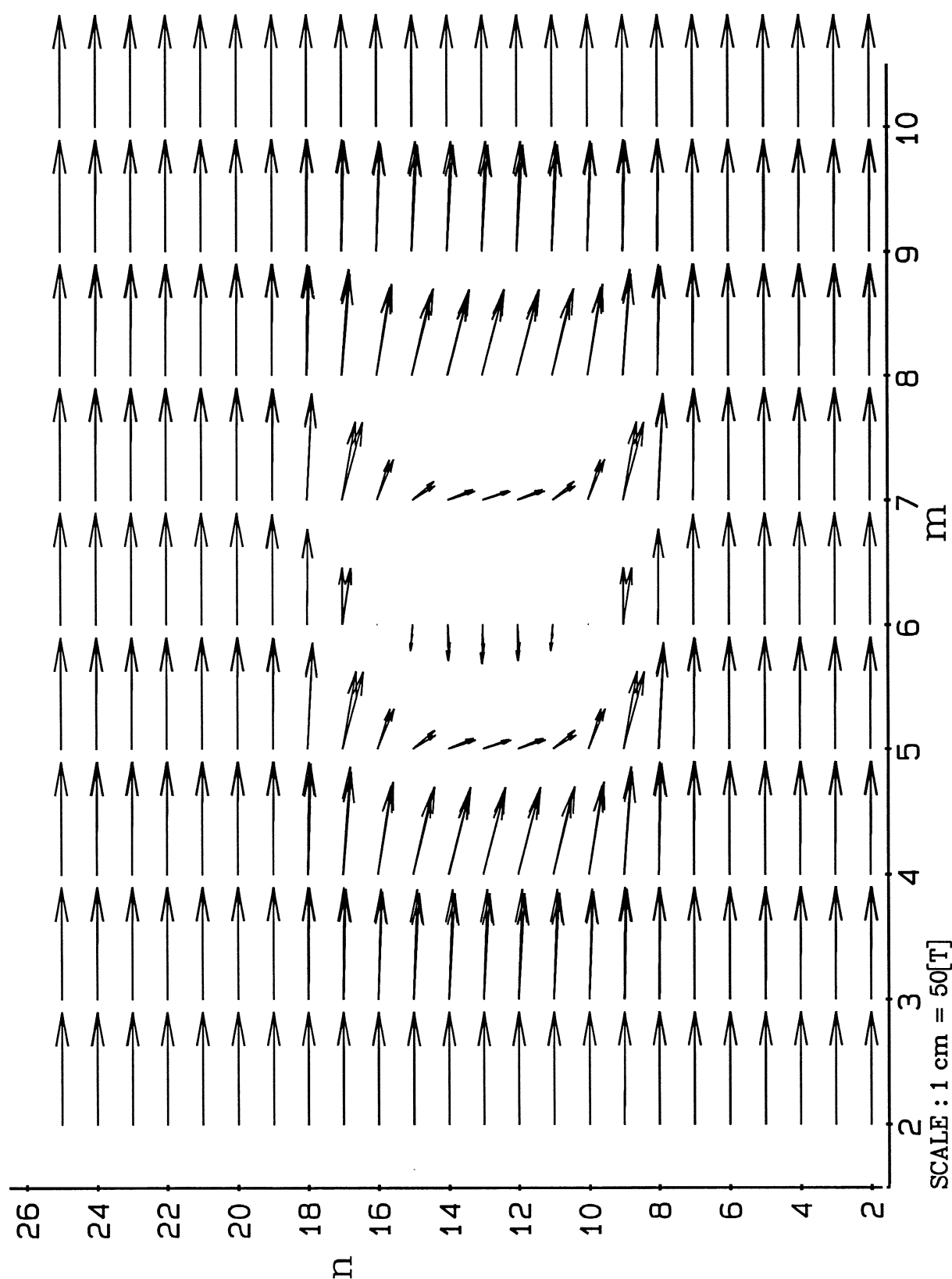
Nov. 1999

Exkl. B., $\sigma_\varphi = 0.1^\circ$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.7a



Analytische test case voor richtingen, 10x25 model
 Met ISTM ge-indentificeerde transportenvelden.
 T-veld van Figuur B.3.3b en T^* -veld van Figuur B.3.6

Non-Linear
 ISTM

Nov. 1999

Exkl. B., $\sigma_\varphi = 0.1^\circ$

DELFT HYDRAULICS

Z2727.30

Fig. B.3.7b